International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 6, December 2026, pp. 4855∼4866
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i6.39593

Assessing learning progress and teaching effectiveness in the era of generative AI: a dual bibliometric analysis and systematic review

Said Maizzou[1], Abdelhak Anajar[1], Youness Ezzaaime[2], El Amine Rouaga[2]
[1]Laboratory LASMO, National School of Business and Management, Hassan First University, Settat, Morocco
[2]Laboratory LARETA, Faculty of Economics and Management, Hassan First University, Settat, Morocco

Article Info

Article history:

Received Mar 18, 2026
Revised Aug 6, 2026
Accepted Sep 30, 2026

Keywords:

Bibliometric analysis
Generative AI
Learning assessment
Systematic literature review
Teaching effectiveness

ABSTRACT

Despite the rapid integration of generative artificial intelligence (GenAI) into higher education, educators and policymakers face a critical challenge: the absence of a comprehensive synthesis identifying how GenAI tools affect the evaluation of learning progress and teaching effectiveness. Existing studies remain fragmented and lack systematic mapping of research trends, dominant methodologies, and research gaps. To address this, the present study adopts a dual methodological approach combining bibliometric analysis and a systematic literature review. A bibliometric analysis of 363 Scopus-indexed publications (2015–2025) was conducted using VOSviewer to map publication trends, influential authors, productive countries, and dominant themes. This was complemented by a systematic review of 35 high-quality empirical studies. Findings reveal an exponential surge in publications post-2022, driven by the democratization of large language models (LLMs) such as ChatGPT. GenAI enhances personalized feedback, self-regulated learning (SRL), and formative assessment at scale, while raising substantive concerns about academic integrity, algorithmic bias, and the erosion of outcome-based evaluation frameworks. The study concludes that sustainable GenAI integration requires human-centered pedagogical designs, institutional governance, and equity-focused policy, offering a structured evidence base for researchers, educators, and policymakers.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Said Maizzou
Laboratory LASMO, National School of Business and Management, Hassan First University
Settat, Morocco
Email: said.maizzou.doc@uhp.ac.ma

1. INTRODUCTION

The recent development of generative artificial intelligence (GenAI) and large language models (LLMs) has radically altered the process of teaching and learning, opening new opportunities for personalisation, assessment, and educational feedback. The application of artificial intelligence (AI), in the form of ChatGPT and other LLMs, is gaining increasing scientific attention, and one of its central questions is: how do we measure the progress of student learning and the effectiveness of teaching when GenAI is integrated into educational practice? This question has generated a rapidly growing literature over recent years [1]–[3].

This growth has not yet brought cohesion but the field is still divided. Researchers, educators, and policymakers are presented with a tangible challenge that lacks a systematic pattern of development or a map of how the field of GenAI-supported assessment of learning and teaching (AL&T) has evolved, which themes are emerging, which ones are missing, and what evidence-based guidance might be inferred for practice and policy. Existing research focuses on one aspect at a time, e.g. personalized feedback, academic integrity, and/or adoption factors, and often lacks the integrated overview needed by practitioners to inform pedagogical and institutional choices. This study is prompted directly by this.

There are several empirical studies that show LLM can support teachers' design of teaching situations, simulate the teaching experience and enhance the lesson plan, which can help to increase teachers' perceived effectiveness and adapt content to learners' needs [1], [4]. There are other studies that examine the use of ChatGPT in higher education, and that identify factors linked to perceived usefulness, ease of use, and academic social norms [2], [5]. There is consistent evidence from systematic reviews that GenAI tools can foster autonomous learning, cognitive engagement, and the development of higher-order skills, but they can also pose challenges to academic honesty and assessment practices [3], [6]. In a similar vein, multiple meta-analyses have tried to measure the effects of ChatGPT on academic performance, revealing mixed effects across different discipline settings and usage patterns of the tool [7], [8].

It is no longer possible to think of the assessment of learning in the context of traditional outcome-based approaches. It also needs to be robust enough to address processes like co-authoring documents, problem solving, and reasoning with support, which blur the distinction between learning support and cognitive delegation [9], [10]. Concurrently, the effectiveness of teaching is also judged by teachers' critical and responsible application of these tools in their teaching [11]. The need to create novel assessment approaches that take a holistic perspective of the impact of GenAI in education through combining quantitative indicators (achievement, learning progress) with qualitative indicators (motivation, autonomy, well-being, and perception) is thus emphasized in the literature [8], [12].

In this context, this study makes two methodological contributions: the combination of bibliometric analysis and a systematic literature review, which, to our knowledge, has not been used before in this particular intersection of GenAI and educational evaluation. The bibliometric dimension depicts the intellectual structure and the development of the field at scale, whereas the systematic review gives a detailed and quality-filtered summary of empirical results. Combined, these two lenses provide depth and breadth to uncover: research trends, the dominant themes, the influential actors, and the missing gaps that either lens alone would not reveal.

The study aims to answer the following four research questions which, in turn, frame the bibliometric analysis and systematic review:

- How has the number of scientific publications on the assessment of learning progress and teaching effectiveness using GenAI evolved from 2018 to 2025? (RQ1)
- What are the most cited and published scientific journals and authors for this research area? (RQ2)
- What are the most prolific and influential countries and academic institutions on research on GenAI-assisted learning assessment? (RQ3)
- What are the most common keywords and dominant research themes in the studies which aimed to evaluate learning progress and teaching effectiveness with GenAI? (RQ4)

In this study, questions are approached from two angles, through the bibliometric and systematic content analysis, which will reveal the intellectual structure of the field, highlight emerging research fronts, expose gaps that are not yet explored, and provide an evidence-based basis for future research and policy-making in AI-assisted education.

2. METHOD

This article aims to carry out a systematic literature review of the research literature on the application of GenAI for learning assessment and pedagogical effectiveness from an integrative perspective, based on bibliometric analysis. The aim is to define trends in publications, influential authors and journals, topic and method specific research focus and problems, and future avenues of research. A literature search was performed in Scopus (on 7 January 2026) for peer-reviewed journal papers and conference papers. The following is the exact query that was used to identify publications in the area of intersection of learning/teaching evaluation and GenAI technologies, Scopus search query: ALL (”evaluating learning progress” AND ”teaching

success" AND "generative AI") AND PUBYEAR > 2014 AND PUBYEAR < 2026 AND (LIMIT-TO(DOCTYPE,"ar") OR LIMIT-TO(DOCTYPE,"cp")) AND (LIMIT-TO(LANGUAGE,"English"))

An initial set of 840 documents were returned in response to this query. The following pre-established inclusion and exclusion criteria were then used to screen documents. Inclusion criteria: i) publication year: 2015-2025; ii) document type: peered-reviewed journal articles and conference papers; iii) language: English; and iv) relevance to the context of GenAI or LLMs, and to the evaluation of learning progress and teaching effectiveness. Exclusion criteria: i) editorial papers, book chapters and grey literature; ii) papers that are not directly related to educational assessment or pedagogical effectiveness; and iii) duplicates. Following systematic screening, 363 publications remained to be analyzed by means of bibliometric analysis, of which 35 high quality empirical studies were selected for systematic content analysis. Selection process was conducted according to the guideline of preferred reporting items for systematic reviews and meta-analyses (PRISMA), as shown in Figure 1. The PRISMA flow diagram outlines the process of identification, screening, eligibility, and final inclusion, ensuring transparency, rigor, and repeatability in the review process.

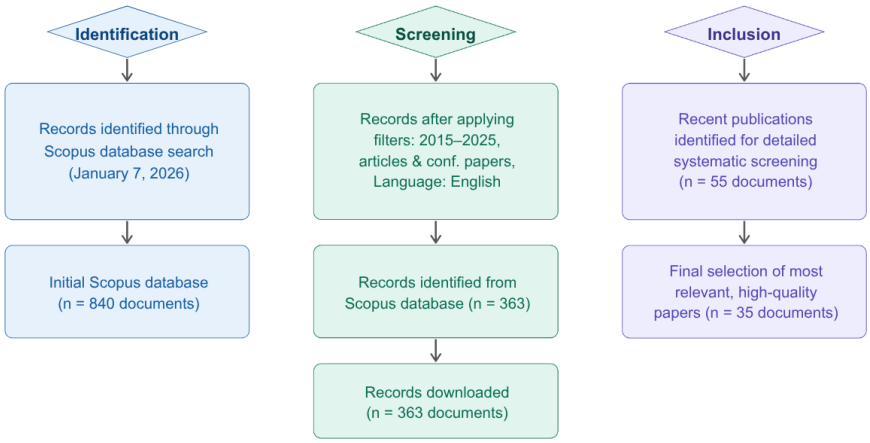


Figure 1. PRISMA diagram summarizing the study screening and selection stages

For the bibliometric analysis, VOSviewer (version 1.6.20) was used to extract and visualize metadata from the 363 retained publications. Analyses covered: co-authorship networks (country and institution level), co-occurrence networks of author keywords (minimum co-occurrence threshold: 5), and source clustering. For the systematic review, a structured coding protocol was applied to the 35 selected studies. Each study was coded along the following dimensions: i) country and institutional affiliation; ii) educational level and disciplinary context; iii) research methodology (experimental, quasi-experimental, qualitative, mixed-methods); iv) type of GenAI tool used; v) learning outcomes measured; vi) key findings; and vii) ethical or institutional concerns reported. Coding was conducted independently by two of the three authors; disagreements were resolved through discussion. Inter-rater reliability was assessed using Cohen's kappa ($\kappa = .81$), indicating strong agreement.

3. RESULTS

A bibliometric analysis of Scopus-indexed publications was conducted to examine research on learning assessment and teaching effectiveness using GenAI. Publication trends, influential authors and journals, international collaborations, and keyword patterns were analyzed to identify key themes, research gaps, and high-impact work guiding future studies.

3.1. Research question 1: how has the number of scientific publications on the assessment of learning progress and teaching effectiveness using GenAI evolved from 2018 to 2025?

To answer this first question, an analysis of the distribution of publications by year was carried out for the period from 2018 to 2025. The graph in Figure 2 illustrates the evolution of the number of scientific publications devoted to GenAI in the AL&T effectiveness over the period studied. Several major lessons can be drawn from this analysis. The overall trend shows an exponential increase in the amount of publications. The

scientific output remained rather low till 2023 (27 documents or less) and then showed spectacular growth from 2023 onwards. From 27 documents in 2023 to 212 documents in 2025, the number of publications (TP) has exploded, demonstrating the increased scientific interest in the integration of GenAI in educational assessment processes within just a few years.

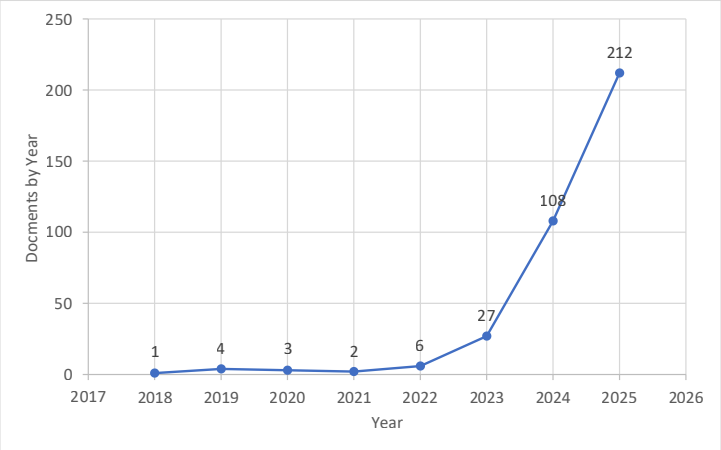


Figure 2. Temporal evolution of scientific publications on GenAI in learning assessment (2018–2025)

Such a speed-up is not accidental. The post-2022 surge aligns perfectly with the public launch and unprecedented democratisation of large language model–based tools, specifically by OpenAI with its release of ChatGPT in November 2022. After this event, there was a proliferation of empirical studies in educational contexts globally: researchers and practitioners tried to evaluate the impact of these on learning, teaching and assessment. The policy interest, both at institutional and governmental levels, including the national AI education policies and UNESCO recommendations on AI in education, also contributed to the increase in publications, creating demand for evidence-based policies. This is because technological accessibility coupled with policy need has been the reason behind the dramatic increase in research that has been seen since 2023. The graph also shows that 2025 is likely to be a peak year for production (212 documents) and that production is likely to moderate in 2026 (around 250 documents predicted), possibly marking an early consolidation phase in fields.

3.2. Research question 2: what are the most cited and published scientific journals and authors for this research area?

A bibliometric analysis was performed to determine the most relevant scientific journals and authors in this field in order to answer this second research question. The TP, the total number of citations (TC), the CiteScore of the journal, the most cited article and its TC were used for analysis of the most cited journals, as shown in Table 1. Table 1 shows the most productive and influential journals and conferences related to the use of GenAI for educational assessment. The analysis shows that the publishing market is quite varied. On the other hand, the absolute publication number (TP $>$ 40,000) and TC are dominated by large-scope multidisciplinary mega-journals (such as Sustainability, IEEE Access, and Applied Sciences) which illustrate the interdisciplinary interest in advanced technologies. On the other hand, more specialized titles stand out for their targeted impact and thematic relevance. The CHI Conference and the International Journal of Human-Computer Interaction have high CiteScores, highlighting the importance of human-computer interaction in this field. Most significantly, Education and Information Technologies has the highest CiteScore among the sources listed in Table 1 (11.8). CiteScore measures the average number of citations received per document over a four-year period in Scopus; therefore, a value of 11.8 indicates a comparatively high citation impact for this journal within the set of sources analyzed. Regarding the most prolific authors in the field, the analysis criteria used were: author name, year of first publication, total TP, h-index, TC, current institutional affiliation, and country, as presented in Table 2.

Table 1. Leading journals and conferences by publication volume and citation impact

Journal/Conference	TP	TC	CS	Most cited article	TC*	Pub.
CHI Conference Proceedings	5,054	49,182	9.7	Tangible Bits: Towards Seamless...	2700	ACM
Frontiers in Education	3,670	13,418	3.7	A Critical Review of Student Self...	381	Front.
IEEE Access	44,988	402,859	9.0	Millimeter Wave Mobile for 5G	6612	IEEE
Int. Journal of HCI	1,095	9,552	8.7	Evaluation of System Usability Scale	3758	T&F
Applied Sciences	49,646	275,395	5.5	ML for Intrusion Detection	877	MDPI
ACL Proceedings	5,956	45,765	7.7	BLEU: Method for Evaluation of MT	22402	ACL
TMLR	1,785	4,762	2.7	Emergent Abilities of LLMs	899	TMLR
Education and IT	2,440	28,739	11.8	VR Head-Mounted Displaysin Ed.	1140	Spring.
Electronics	17,336	105,744	6.2	ML Interpretability Survey	1418	MDPI
Sustainability	58,414	450,030	7.7	Restoring Soil Quality	1410	MDPI

Note: CS = CiteScore; TC* = times cited of the most cited article. Source: Scopus (2025).

Table 2. Most influential authors by publication volume and citation impact

Author	Year*	TP	h-index	TC	Current affiliation	Country
Gašević, D.	2003	693	76	23,238	Monash University	Australia
Leinonen, J.	2015	137	27	3,399	Aalto University	Finland
Prather, J.	2017	72	23	2,010	Baylor University	USA
Becker, B. A.	2011	96	26	2,847	University Coll. Dublin	Ireland
Denny, P.	2009	196	35	5,127	Univ. of Auckland	New Zealand

Note: Year* = year of first Scopus-indexed publication. Source: Scopus (2025).

3.3. Research question 3: what are the most prolific and influential countries and academic institutions on research on GenAI-assisted learning assessment?

Figure 3 generated using VOSviewer software, highlights the networks of scientific collaboration between countries. The United States (U.S.) occupies the most central and predominant position in the network, leading a major cluster that includes close collaborations with Ireland, Finland, the Czech Republic, and Estonia. A second major hub, China, shows a strong structure with Asian and Pacific partners such as Hong Kong, Taiwan, Indonesia, Pakistan, and Malaysia. The United Kingdom seems to be a key link between these centers, closely linked with the U.S. and also continental Europe. There is also a marked European cluster under the leadership of Germany and Italy. Centrality is still focused on countries with a powerful technological and research base, with the development of countries like India, Iran and South Africa marking a growth in the network towards the Global South. In conclusion, the findings suggest that the scientific output of GenAI-assisted learning assessment is concentrated in a small number of countries with substantial research capacities and that these results demonstrate a fairly slow geographical, albeit developing, diversification, with increasing interest in this research field globally.

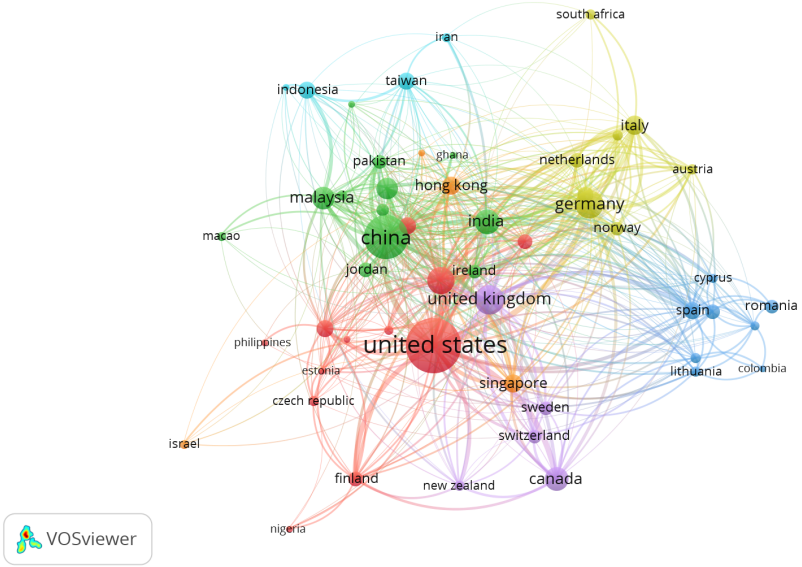


Figure 3. Network of scientific collaboration between countries in GenAI educational research

3.4. Research question 4: what are the most common keywords and dominant research themes in the studies which aimed to evaluate learning progress and teaching effectiveness with GenAI?

A co-occurrence analysis of 1,253 keywords (distinct authors) has been done. The keyword co-occurrence map in Figure 4 shows that there are three interrelated thematic clusters. At the centre of the corpus are the core technologies, which are conceptualized in the first cluster: AI, ChatGPT, GenAI, and LLMs. The second cluster includes assessment and learning outcomes: personalized learning, student engagement, and academic performance. The third group represents human, ethical and institutional factors, challenges, technology adoption, and self-efficacy. The three clusters in total outline the intellectual framework of the field and suggest directions for future research for the next few years.

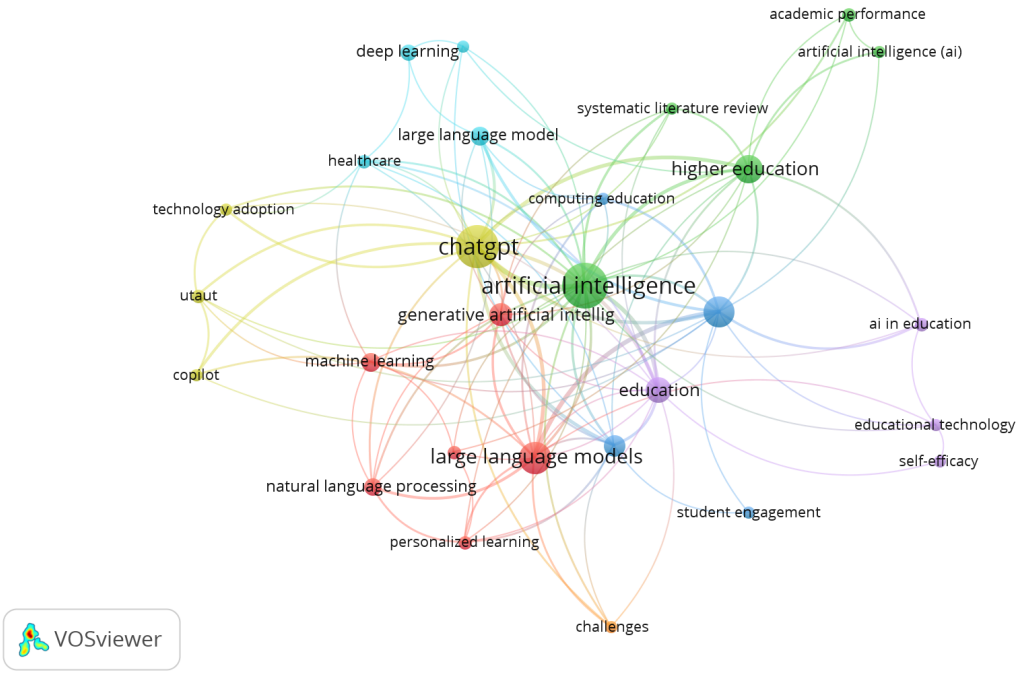


Figure 4. Co-occurrence network of author keywords in GenAI for educational assessment

4. SYSTEMATIC REVIEW

The systematic review of 35 selected studies offers a comprehensive understanding of current research investigating the application of GenAI, such as ChatGPT and other LLMs, to measure learning progress and teaching effectiveness. Selected studies cover the timeframe from 2015 to 2025 and show trends in the use of pedagogical practice, assessment method, learning outcomes and ethical issues.

4.1. Applications of GenAI in education

The use of GenAI has expanded to various learning environments, including higher education, secondary education, vocational education, and language education. Previous studies have demonstrated that ChatGPT can support coding education through individualised tutoring, feedback, and assessment assistance, with students developing higher-order thinking skills [7], [13]. In language learning, AI-driven conversational agents have been shown to enhance empathy and oral communication skills, while recent evidence suggests that ChatGPT can outperforms human instructors in English essay grading, although its accuracy appears lower in other languages [14], [15]. GenAI is also being implemented as a direct mediator in the teaching process within substitution, augmentation, modification, and redefinition (SAMR) framework models. Furthermore, immersive technologies such as the metaverse, when coupled with GenAI, present new paradigms of collaborative learning and human-computer interaction [4], [11].

4.2. Impact on learning outcomes and student engagement

Multiple studies examined both cognitive and motivational impacts of AI-supported learning. Individualized feedback and adaptive instruction were found to enhance engagement [16]. AI interventions can foster knowledge acquisition through systematic learning, improve intrinsic motivation, and support knowledge integration [17]–[19]. The combination of AI with social and collaborative systems has also been demonstrated to support student mental health and well-being and promote smart learning environments [8].

However, a few studies have been released that are concerned about potential adverse effects. AI can make it look like someone has a better understanding than they actually do, which can lead to a false sense of mastery in the beginning, potentially reducing long-term retention of knowledge when relying on AI passively. The results underscore the need for scaffolding and dialogue in the use of AI in learning [20].

4.3. Pedagogical and instructional frameworks

Many studies have introduced AI as a tool to enhance teaching and learning. A multi-dimensional model has been proposed in which AI supports the pastoral, cognitive, and instructional aspects of teaching and learning, while also assisting assessment and reflective practice [21]. In addition, Bloom's taxonomy has been mapped to AI tools to support lifelong learning and skill development [22].

Another research studies the relationship between the use of AI and interactive, constructive, active, passive (ICAP) to understand the effects of active AI interaction with learning outcomes [23]. In addition, teacher development has been tackled by the co-creation of learning material with the support of AI, formative assessment and adaptive feedback [4], [15]. There are proposals for hybrid intelligence systems of the future in which humans and AI support one another in learning [24].

4.4. Determinants of AI adoption in education

There are several factors mentioned in the literature that have an impact on AI adoption. Some of the most important predictors are perceived usefulness, ease of use, and self-efficacy [25], [26]. Adoption is also influenced by institutional and contextual factors, such as governance, ethical standards and policy frameworks [27]. University concerns about academic integrity, algorithmic bias, and data privacy may be adoption barriers, although the perception of gains in efficiency outweighs these concerns in most cases [28], [29].

4.5. Ethical, social, and cognitive considerations

The responsible use of AI in education involves ethical and sociotechnical aspects. Researchers highlight the importance of equality of access, monitoring to avoid bias, and protection from the commercialization of learning [30]. AI literacy for educators is crucial to equip learners to reap the benefits of personalisation without becoming overly dependent on AI [22]. AI may also be seen to encourage creativity and thinking skills [31] but the socio-emotional and reflective components of the teacher's role cannot be fully replicated by AI, leaving a need for ongoing human monitoring [24], [32].

4.6. Emerging trends and research directions

There are several emerging trends in the literature identified by the systematic review. The use of GenAI in combination with immersive and collaborative learning spaces, such as VR, AR and metaverse platforms, is being increasingly researched for personalized and collaborative large-scale learning [11]. Pedagogical assessment frameworks are being created with AI to assist both in formative and summative evaluations, without sacrificing pedagogical judgement and ethics [15]. AI's use is spreading beyond higher education into vocational, secondary and professional education [22], [33]. Lastly, equity and student well-being and mental health are emerging as themed priorities [8], [34].

5. DISCUSSION

Based on 363 publications from Scopus (bibliometric analysis) and 35 empirical studies from the systematic review, the results of this study show that this area has been structured and transformed in a very short time since the democratisation of LLM in late 2022. The next subsections offer a critical interpretation of these findings with regard to five dimensions: pedagogical implications, cognitive consideration, socio-emotional consideration, determinants of adoption, ethical and societal implications, and gaps in the research.

5.1. Pedagogical implications

AI's use in education shows great promise for transforming personalised learning. It has been found that the use of AI feedback and interactivity contributes to SRL, increases students' motivation, and facilitates the development of higher-order thinking [7], [17], [18]. AI can also aid teachers by automating repetitive grading tasks and providing detailed analysis, so that educators can focus on the most crucial aspects of their work, including strategic and socio-emotional learning [4], [15].

However, these benefits need to be taken with a grain of salt. The use of AI tools for passive engagement could lead to a false sense of competence and diminished long-term retention if students do not critically interact with the content. These are especially problematic in first year or competency courses where basic level of mastery is a prerequisite. Rather than an outright ban on AI, the pedagogical implication is that its integration needs to be integrated into a structured, scaffolding framework (e.g., SAMR, Bloom's taxonomy, and ICAP) where active, constructive engagement is fostered instead of passive consumption. This presents a challenge for educators to align their learning sequences with the use of AI, enabling students to reflect, evaluate and use AI-generated content.

5.2. Cognitive, social, and emotional considerations

The cognitive ability, creativeness and analytical thinking are capabilities that GenAI has proven useful in building up [31]. AI tools that can help foster a communal spirit, empathy, and engagement, such as virtual agents and immersive environments, can be used in collaborative applications. AI tools can be used to create collaborative applications, such as virtual agents and immersive environments, that can help foster communication skills, empathy, and engagement [11], [14]. However, AI is not a complete substitute for the social and emotional aspects of learning and teaching, and human intervention is still necessary to make learning environments emotionally, morally and interpersonally responsive [24], [32]. This is a key takeaway: GenAI should not replace but be integrated into human educational relationships. These roles of teachers, such as those of motivation, ethical modeling, and affective responding cannot be replaced by current AI architectures. Education professionals should be therefore provided with the tools to understand how their work can be better complemented by AI tools, while other parts of the process must be in direct human hands.

5.3. Determinants of adoption and implementation

Institutional, technological, and individual factors all play a part in the adoption of AI for educational practice. These four factors of perceived usefulness, ease of use, self-efficacy and institutional support are always found to be positive predictors [25]–[27]. On the other hand, issues related to academic integrity, data privacy and algorithmic bias, suggest the importance of establishing strong governance systems and provisions in policy documents [28], [30].

The results of these studies have practical implications for policy makers. Equitable adoption of AI requires investments in AI literacy programs for educators and students. Data sovereignty and algorithmic transparency need to be addressed through regulatory frameworks, especially in multilingual and culturally diverse education systems, where there may be a bias in the culture of the education system in which AI models were trained. Finally, the bibliometric analysis clearly exposes the concentration of productive research within the realm of the U.S., China, and Northern Europe, which highlights a potential danger of imbalances in the transfer of knowledge and technologies, a danger which policy should be prepared to act to mitigate through collaboration and capacity building in the Global South.

5.4. Ethical and societal implications

There are substantial moral and social issues involved in the use of AI in the education sector. Researchers give recommendations on algorithmic prejudices, over-reliance on automation, and the commercialization of learning [30]. The growth of AI literacy for both teachers and students is essential for ensuring an AI tool's use to support, not replace, human judgment [22].

On the societal level, the rapid integration of GenAI in education has the potential to widen existing inequalities if there is unequal access to high-quality AI tools and their ability to be used for critical purposes. Therefore, institutions and policy makers need to embrace the principles of equity-by-design in procurement, deployment and evaluation of AI. This involves adapting algorithmic bias audits for local language and cultural contexts and does not create an additional disadvantage for historically marginalized student groups in the context of assessment reform based on AI.

5.5. Research gaps, limitations, and future directions

There are still several gaps in evidence that are evident. First, there are no long-range cognitive or SE outcomes studies in the current corpus and many of the studies are short term and superficially experimental. Second, the differential effect of AI by discipline, age, and culture has yet to be explored. Third, there are vague theoretical conceptions of how to implement AI ethically and institutionally in education. Fourth, there is a lack of theoretical development in relation to the implementation of ethical and institutional design in education. Such gaps need to be overcome for sustainable integration to be informed.

There are also some limitations to this study that need to be noted. The bibliometric analysis only included publications in English from Scopus indexed publications, which could underrepresent research from non-Anglophone regions and contexts. The systematic review followed a PRISMA-compliant approach but used a prespecified search string which could not ensure the inclusion of all empirical studies with other wording of the search query. Future studies should focus on longitudinal, cross-cultural designs and use multilingual databases, as well as consider grey literature. It is also important to have instruments for measuring learning outcomes from AI-assisted learning that are validated both cognitively and affectively, supporting more in-depth cumulative synthesis.

6. CONCLUSION

This research work is a thorough dual-method synthesis of research related to the use of GenAI to evaluate learning progress as well as teaching effectiveness. The research indicates that the volume of literature on GenAI has rapidly increased since 2022, as evidenced by both the increasing TP on Scopus and the number of empirical studies conducted, which is steadily rising. The study concludes that there has been an exponential increase in the TP on GenAI since 2022, fuelled by the democratization of LLMs and an increased interest in AI-supported education in policies. Bibliometric mapping reveals the U.S., China, and Northern Europe as the major knowledge hubs with Education and Information Technologies Conference Proceedings and the CHI Conference Proceedings as the most prominent outlets. The results of keyword co-occurrence analysis identify three thematic clusters, representing the core AI technologies, learning assessment processes, and ethical-institutional dimensions, which together form the intellectual architecture of the field.

The findings of the systematic review highlight the potential for meaningful improvements in personalized feedback, SRL, formative assessment at scale, and higher-order thinking with the use of GenAI, as well as significant substantive concerns regarding academic integrity, algorithmic bias, the illusion of competence in novice learners and passive cognitive dependency. The function of teacher remains unchanged, but is reconfigured: AI creates space for teachers to focus on socio-emotional and strategic learning, and demands teachers to learn new skills in AI pedagogy. The results highlight the need for equity-oriented approach to AI policy and practices, multilingual bias auditing, investment in AI literacy, and international partnership to close knowledge and technology transfer gaps between the Global North and South.

Longitudinal, cross-cultural, and cross-disciplinary study designs should be given more priority to address the limitations of the current body of evidence that are context-specific and short-term. There is an urgent need for validated instruments that assess cognitive and affective learning outcomes in environments that are mediated by AI in order to provide for cumulative synthesis. In conclusion, the successful and responsible adoption of GenAI in education requires a balance between technology and human-centered pedagogies, a balance that will require continued interaction and partnership among researchers, educators, institutions and policy makers.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Said Maizzou	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	
Abdelhak Anajar	✓	✓		✓	✓	✓		✓	✓		✓		✓	
Youness Ezzaaime		✓			✓	✓	✓			✓		✓		✓
El Amine Rouaga		✓			✓	✓	✓			✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

ETHICAL APPROVAL

Ethical approval and informed consent were not required.

DATA AVAILABILITY

Data are available from the corresponding author upon reasonable request.

REFERENCES

[1] B. Hu, J. Zhu, Y. Pei, and X. Gu, "Exploring the potential of LLM to enhance teaching plans through teaching simulation," *npj Science of Learning*, vol. 10, no. 1, p. 7, Feb. 2025, doi: 10.1038/s41539-025-00300-x.

[2] I. Dorobăţ and A. M. I. Corbea (Florea), "Assessing ChatGPT adoption in higher education: an empirical analysis," *Electronics*, vol. 14, no. 23, p. 4739, Dec. 2025, doi: 10.3390/electronics14234739.

[3] F. J. Agbo, C. Olivia, G. Oguibe, I. T. Sanusi, and G. Sani, "Computing education using generative artificial intelligence tools: a systematic literature review," *Computers and Education Open*, vol. 9, p. 100266, Dec. 2025, doi: 10.1016/j.caeo.2025.100266.

[4] P. Wang, Y. Jing, and S. Shen, "A systematic literature review on the application of generative artificial intelligence (GAI) in teaching within higher education: instructional contexts, process, and strategies," *The Internet and Higher Education*, vol. 65, p. 100996, Apr. 2025, doi: 10.1016/j.iheduc.2025.100996.

[5] H. M. Al-Hattami, "Empowering business research with ChatGPT: academic and student insights through the UTAUT lens," *Discover Computing*, vol. 28, no. 1, p. 179, Aug. 2025, doi: 10.1007/s10791-025-09692-1.

[6] N. Raihan, M. L. Siddiq, J. C. S. Santos, and M. Zampieri, "Large language models in computer science education: a systematic literature review," in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1* Feb. 2025, pp. 938–944, doi: 10.1145/3641554.3701863.

[7] R. Deng, M. Jiang, X. Yu, Y. Lu, and S. Liu, "Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies," *Computers & Education*, vol. 227, p. 105224, 2025, doi: 10.1016/j.compedu.2024.105224.

[8] M. F. Shahzad, S. Xu, H. Liu, and H. Zahid, "Generative artificial intelligence (ChatGPT-4) and social media impact on academic performance and psychological well-being in China's higher education," *European Journal of Education* vol. 60, no. 1, p. e12835, Mar. 2025, doi: 10.1111/ejed.12835.

[9] S. Bernstein, A. Rahman, N. Sharifi, A. Terbish, and S. MacNeil, "Beyond the benefits: a systematic review of the harms and consequences of generative AI in computing education," in *Proceedings of the 25th Koli Calling International Conference on Computing Education Research*, Nov. 2025, pp. 1–18, doi: 10.1145/3769994.3770036.

[10] R. Adnin, A. Pandkar, B. Yao, D. Wang, and M. Das, "Examining student and teacher perspectives on undisclosed use of generative AI in academic work," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Apr. 2025, pp. 1–17, doi: 10.1145/3706598.3713393.

[11] Y. Wang, L. Wang, and K. L. Siau, "Human-centered interaction in virtual worlds: a new era of generative artificial intelligence and metaverse," *International Journal of Human–Computer Interaction*, vol. 41, no. 2, pp. 1459–1501, Jan. 2025, doi: 10.1080/10447318.2024.2316376.

[12] F. X. Baskara, "ChatGPT-4: a fresh lens on holistic education's reimagination," in *AIP Conference Proceedings*, 2025, p. 020008, doi: 10.1063/5.0261979.

[13] M. B. Garcia, "Teaching and learning computer programming using ChatGPT: a rapid review of literature amid the rise of generative AI technologies," *Education and Information Technologies*, vol. 30, no. 12, pp. 16721–16745, Aug. 2025, doi: 10.1007/s10639-025-13452-5.

[14] A. Derakhshan, T. Teo, and S. Khazaie, "Investigating the usefulness of artificial intelligence-driven robots in developing empathy for English for medical purposes communication: the role-play of Asian and African students," *Computers in Human Behavior*, vol. 162, p. 108416, Jan. 2025, doi: 10.1016/j.chb.2024.108416.

[15] E. Fokides and E. Peristeraki, "Comparing ChatGPT's correction and feedback comments with that of educators in the context of primary students' short essays written in English and Greek," *Education and Information Technologies*, vol. 30, no. 2, pp. 2577–2621, Feb. 2025, doi: 10.1007/s10639-024-12912-8.
[16] H. Yaseen, A. S. Mohammad, N. Ashal, H. Abusaimeh, A. Ali, and A.-A. A. Sharabati, "The impact of adaptive learning technologies, personalized feedback, and interactive AI tools on student engagement: the moderating role of digital literacy," *Sustainability*, vol. 17, no. 3, p. 1133, Jan. 2025, doi: 10.3390/su17031133.
[17] D. T. K. Ng, C. W. Tan, and J. K. L. Leung, "Empowering student self-regulated learning and science education through ChatGPT: a pioneering pilot study," *British Journal of Educational Technology*, vol. 55, no. 4, pp. 1328–1353, Jul. 2024, doi: 10.1111/bjet.13454.
[18] Z. Gao, J.-H. Cheah, X.-J. Lim, and X. Luo, "Enhancing academic performance of business students using generative AI: an interactive-constructive-active-passive (ICAP) self-determination perspective," *International Journal of Management Education*, vol. 22, no. 2, p. 100958, Jul. 2024, doi: 10.1016/j.ijme.2024.100958.
[19] I. S. Chaudhry, S. A. M. Sarwary, G. A. El Refae, and H. Chabchoub, "Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—a case study," *Cogent Education*, vol. 10, no. 1, pp. 1–30, Dec. 2023, doi: 10.1080/2331186X.2023.2210461.
[20] J. Prather *et al.*, "The widening gap: the benefits and harms of generative AI for novice programmers," in *Proceedings of the 2024 ACM Conference on International Computing Education Research - Volume 1*, Aug. 2024, pp. 469–486, doi: 10.1145/3632620.3671116.
[21] H. Yusuf, A. Money, and D. Daylamani-Zad, "Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art," *Educational Technology Research and Development*, vol. 73, no. 2, pp. 815–874, Apr. 2025, doi: 10.1007/s11423-025-10447-4.
[22] A. Abulibdeh, E. Zaidan, and R. Abulibdeh, "Navigating the confluence of artificial intelligence and education for sustainable development in the era of Industry 4.0: challenges, opportunities, and ethical dimensions," *Journal of Cleaner Production*, vol. 437, p. 140527, Jan. 2024, doi: 10.1016/j.jclepro.2023.140527.
[23] J. Liang, L. Wang, J. Luo, Y. Yan, and C. Fan, "The relationship between student interaction with generative artificial intelligence and learning achievement: serial mediating roles of self-efficacy and cognitive engagement," *Frontiers in Psychology*, vol. 14, p. 1285392, Dec. 2023, doi: 10.3389/fpsyg.2023.1285392.
[24] D. Dellermann, A. Calma, N. Lipusch, T. Weber, S. Weigel, and P. Ebel, "The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems," in *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*, 2019, pp. 1–10, doi: 10.24251/HICSS.2019.034.
[25] C.-J. Fu, A. D. K. Silalahi, S.-C. Huang, D. T. T. Phuong, I. J. Eunike, and Z.-H. Yu, "The (Un)knowledgeable, the (Un)skilled? Undertaking Chat-GPT users' benefit-risk-coping paradox in higher education focusing on an integrated, UTAUT and PMT," *International Journal of Human–Computer Interaction*, vol. 41, no. 9, pp. 5663–5693, May 2025, doi: 10.1080/10447318.2024.2365028.
[26] C. Y. Lai, K. Y. Cheung, and C. S. Chan, "Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: an extension of the technology acceptance model," *Computers and Education: Artificial Intelligence*, vol. 5, p. 100178, 2023, doi: 10.1016/j.caeai.2023.100178.
[27] A. Erdmann and L. Toro-Dupouy, "The influence of the institutional environment on AI adoption in universities: identifying value drivers and necessary conditions," *European Journal of Innovation Management*, vol. 28, no. 9, pp. 4365–4398, Oct. 2025, doi: 10.1108/EJIM-04-2024-0407.
[28] S. A. Bin-Nashwan, M. Sadallah, and M. Bouteraa, "Use of ChatGPT in academia: academic integrity hangs in the balance," *Technology in Society*, vol. 75, p. 102370, Nov. 2023, doi: 10.1016/j.techsoc.2023.102370.
[29] Y. Wang, "When artificial intelligence meets educational leaders' data-informed decision-making: a cautionary tale," *Studies in Educational Evaluation*, p. 100872, Jun. 2021, doi: 10.1016/j.stueduc.2020.100872.
[30] I. O. Gallegos *et al.*, "Bias and fairness in large language models: a survey," *Computational Linguistics*, vol. 50, no. 3, pp. 1097–1179, Sep. 2024, doi: 10.1162/coli_a_00524.
[31] S. Krakowski, "Human-AI agency in the age of generative AI," *Information and Organization*, vol. 35, no. 1, p. 100560, Mar. 2025, doi: 10.1016/j.infoandorg.2025.100560.
[32] M. Sallam, N. Salim, M. Barakat, and A. Al-Tammemi, "ChatGPT applications in medical, dental, pharmacy, and public health education: a descriptive study highlighting the advantages and limitations," *Narra J*, vol. 3, no. 1, p. e103, Mar. 2023, doi: 10.52225/narra.v3i1.103.
[33] H. Jo and D.-H. Park, "AI in the workplace: examining the effects of ChatGPT on information support and knowledge acquisition," *International Journal of Human–Computer Interaction*, vol. 40, no. 23, pp. 8091–8106, Dec. 2024, doi: 10.1080/10447318.2023.2278283.
[34] A. ur Rehman, R. K. Behera, M. S. Islam, F. A. Abbasi, and A. Imtiaz, "Assessing the usage of ChatGPT on life satisfaction among higher education students: the moderating role of subjective health," *Technology in Society*, vol. 78, p. 102655, Sep. 2024, doi: 10.1016/j.techsoc.2024.102655.

BIOGRAPHIES OF AUTHORS



Said Maizzou is a doctoral researcher at the National School of Commerce and Management, affiliated with the Laboratory of Strategy and Organization Management (LASMO) at Hassan First University. His research lies at the intersection of public policy evaluation, labor economics, development studies, and artificial intelligence, with a particular emphasis on youth employability, socio-professional integration, and the effectiveness of state intervention programs in emerging economies. He specializes in impact evaluation methodologies for public programs related to the socio-professional integration of young people and young graduates in Morocco, as well as the application of artificial intelligence in education, public policy, and socio-economic analysis. He can be contacted at email: said.maizzou.doc@uhp.ac.ma.



Abdelhak Anajar is a professor, researcher, and academic affiliated with the National School of Commerce and Management and the Faculty of Sciences and Techniques, both part of Hassan First University. He is also a member of the Laboratory of Strategy and Organization Management (LASMO). His research interests focus on management sciences, organizational performance, public policy analysis, strategic management, youth employability, and the socio-professional integration of young graduates in Morocco. He is particularly recognized for his expertise in youth employment policies and professional integration programs for graduates in the Moroccan context. He can be contacted at email: abdelhak.anajar@uhp.ac.ma.



Youness Ezzaaime is a doctoral researcher in Economics and Management Sciences at Hassan First University. He is also an administrative executive and a Ph.D. student specializing in labor market studies and employment dynamics. His research interests include labor market dynamics, youth employment, higher education outcomes, employability, and public policy evaluation, with a particular focus on the integration of graduates into the labor market and workforce development issues in Morocco. He can be contacted at email: y.ezzaaime@uhp.ac.ma.



El Amine Rouaga is a doctoral researcher in Theoretical and Applied Economics at Hassan First University. He holds a bachelor's degree in Strategic Management of Organizations from Moulay Ismail University and a master's degree in Governance, Human Resource Management, and Skills Engineering from Mohammed V University. His research interests focus on renewable energy, artificial intelligence, numerical analysis, economic modeling, and strategic management. His academic work also explores the application of AI and quantitative methods to organizational performance, sustainable development, and economic decision-making. He can be contacted at email: e.rouaga.doc@uhp.ac.ma.