Academic Integrity in the Age of Generative AI: Challenges to Critical Thinking and Ethical Scholarship

Dr. Muhammad Anees ul Husnain Shah
Associate Professor, Department of Education, University of Education Lahore, D.G Khan Campus, draneesulhusnain@ue.edu.pk

Hafiz Muhammad Rizwan Khan
M.Phil. Education, IER, University of the Punjab, Lahore, Pakistan
rizwanc1984@gmail.com

Dr. Sabir Hussain
Department of Educational Training, the Islamia University of Bahawalpur
sabirjanmarri@gmail.com

Abstract

The dynamic development of generative artificial intelligence (AI) has greatly changed the academic culture in higher education, both bringing some opportunities and challenges to the area of learning, assessment, and ethical scholarship. Although AI applications can be used to improve efficiency and facilitate engagement among students, the proliferation of AI tools has brought up serious questions of academic honesty and the loss of critical thinking abilities. This study critically examines recent empirical research, institutional policies, and international guidelines published between 2023 and 2026 to explore how generative AI influences student cognition, assessment validity, and integrity enforcement. The evidence suggests that unorganized use of AI can cause cognitive offloading, decrease long-term learning, and become more difficult to check authorship, and the traditional models of assessment become more susceptible to AI-generated submissions. Nevertheless, it also indicates that in case of an integrated approach to AI, it could also serve as a scaffold that facilitates reflective thinking and skill acquisition. The study concludes that sustaining academic integrity in the generative AI era requires assessment redesign, transparent disclosure norms, and comprehensive AI literacy education to ensure that critical thinking remains central to higher education.

Keywords: Generative artificial intelligence, academic integrity, critical thinking, higher education, assessment redesign, ethical scholarship

Introduction

The emergence and widespread dissemination of generative AI (GenAI) technologies have posed an organizational problem to the education sector: the very systems that can be used to facilitate learning (e.g., drafting, feedback, explanation) also can replace the very intellectual labor that assessment is supposed to be used to quantify. Academic integrity, in its turn, has ceased to be an issue of cheating or not cheating, but a policy issue--the way institutions formulate a conception of acceptable human-AI cooperation, safeguard the credibility of credentials, and still have assessment demonstrate an expression of high-level thinking (Coates et al., 2025). According to recent studies, policy formulation is underway in universities, but strategies differ in different regions and colleges and still lack uniformity, clarity, and implementation (Jin et al., 2025; Oh & Sanfiljippo, 2025).

GenAI at the cognitive level aggravates the anxiety about the practice of writing without thinking, in which a fluent performance can be a sign of superficial interest, and undermines the practice of analysis, synthesis, and evaluation that higher education is purportedly developing (UNESCO, 2023). It is not merely an issue of student-conduct, but rather a design issue. In the event that evaluations have a strong incentive towards finished goods, GenAI may shortcut or speed up the learning process and thus it may be difficult to distinguish legitimate assistance or unacceptable substitution. As

a result, teachers and schools are moving towards process, reasoning, and accountability type of assessments and clear disclosure and attribution norms (Jin et al., 2025; UNESCO, 2023).

New approaches stress that successful integrity reactions should be systematic instead of being reactive, that is, a combination of AI proficiency, clear policy guidelines, focus on student motive, and regular reporting procedures (Peterson, 2025). Similar parallel initiatives in research governance and scholarly communication also highlight transparency and human responsibility (e.g. disclosure and authorship boundaries) as minimum demands of responsible AI use (Hsu et al., 2025; Smith et al., 2025; Hussain, 2023; Hussain & Abbas, 2023; Hussain & Khoso, 2021; Hussain & Khoso, 2022; Hussain et al., 2024; Perveen & Hussain, 2023; Sheeza et al., 2024; Zafar et al., 2024; Hussain, 2024). Combined, the presented literature indicates that the present moment should not be seen as a transient disturbance, but as a permanent change that necessitates institutions to rebuild the principles of integrity and the framework of evaluation techniques that enable critical thinking to be apparent, provable, and appreciated.

Statement of the Problem

The quick adoption of generative artificial intelligence (AI) in the academic setting has posed an inherent conflict between academic development and upholding academic integrity. Although artificial intelligence technologies can help to improve the efficiency and accessibility of learning, there is a risk that the application of AI-based devices will reduce the ability to think critically, develop original works, and independently study the problem. There is also a rising capacity of students to produce high-quality work with less mental energy involved so that it is hard for teachers to discern between genuine learning and artificially assisted work. The current institutional policies are still unequal and the evaluation models are often unprepared to cope with the substitution of intellectual work that is AI-enabled. Due to this, there is an increasingly acute crisis in higher education: to preserve ethical scholarship and meaningful cognitive development in a world where generative AI can copy, speed up and even substitute fundamental academic work.

Objectives of the Study

1. To examine the impact of generative artificial intelligence on academic integrity and the development of critical thinking skills in higher education.
2. To propose effective policy and pedagogical strategies that promote responsible AI use while preserving ethical scholarship and authentic learning.

Methodology

The present research design is qualitative that is developed basing on an in-depth review and critical analysis of recent academic literature, institutional, and international policies regarding the application of generative artificial intelligence (AI) in academic integrity and critical thinking in higher education. Systematic review of the peer-reviewed journal articles, policy reports, and official frameworks published in 2023-2026 were searched to find out the new themes, challenges, and responses of the institutions.

The thematic analysis strategy was used to group findings on critical areas, such as ethical concerns, assessment redesign, AI governance, and cognitive implications. These themes are synthesized in the study to assess the level at which the generative AI can affect academic practices and intellectual growth. The research is expected to offer a systematic view of the crisis of critical thinking and offer evidence-based suggestions about how AI can be introduced responsibly into the education sector by combining modern scholarly discussion with the analysis of the policy.

Thematic analysis (2023-2026)

Theme 1: Cognitive offloading and weaker durable learning

It has been experimentally observed that, as a shortcut (as opposed to a scaffold), GenAI can lessen the effortful processing required to learn in the long term. In a randomized controlled trial (n=120), students who studied with ChatGPT as an aide scored much lower on a surprise retention exam 45 days later (57.5) than their traditional studying counterparts (68.5) and the argument of cognitive offloading and desirable difficulties holds true (Barcaui, 2025). To supplement this, an experimental mixed-methods study in Ghana (pretest-posttest control group; n = 125) has discovered that the inclusion of ChatGPT has caused quantifiable changes in the dimensions of critical, reflective, and creative thinking, as well as they have shown that cognitive performance is highly contingent on the integration of GenAI into learning activities (Essel et al., 2023). All these studies indicate that GenAI might enhance the performance at the surface and undermine the capacity to build lasting learning, unless the learning experience compels students to assess and justify and reflect on AI outputs (Barcaui, 2025; Essel et al., 2023).

Theme 2: Real-world assessment vulnerability (GenAI can "pass" authentic exams)

One especially influential practical research was done on the possibility of AI penetrating a real university examination system. In one case study of a blind "Turing test" researchers injected fully AI-written submissions into five undergraduate modules and discovered that 94 percent remained undetected and that AI work had a higher grade by approximately half a grade boundary than submissions by actual students (Scarfe et al., 2024). The discovery changes the discourse of hypothetical risk to functional fact: product-based evaluations prove to be a weak signifier of authorship and learning when GenAI will be able to make submissions that are plausible on a scale (Scarfe et al., 2024).

Theme 3: Integrity is shifting from plagiarism to provenance and disclosure and compliance is low

New studies are happily outlining academic integrity as a provenance problem (what role AI played and whether it is recognized). But disclosure regimes are not always practised. When a study of assessment declarations examined the King Business School, 74% of students failed to declare the usage of AI even though they were required to do so; the paper employs the Theory of Planned Behavior to understand how the students failed to do so through the prism of norms, attitudes, and perceptions of behavioral control (Gonsalves, 2025). This implies that policy is not enough, institutions should have a consistent messaging, enforcements that are credible, and designs of assessments that disclosure is viable and significant (Gonsalves, 2025).

Theme 4: Detection tools produce false positives and false negatives (high ethical risk in enforcement)

Evidence of the use of AI-detection strategies experimentally indicates that there exist significant overlaps between human and AI text distribution, and both false positives and false negatives are realistic. According to one of the studies, literature sections were especially likely to be flagged, and it was noted that false positives can be especially detrimental (e.g., false accusation), so people should be careful in considering detectors to be conclusive evidence (Meena and Suriya, 2023). It should be noted that this is part of an increasing trend to suggest that integrity enforcement must be based more on assessment validity, process evidence, and transparent rules rather than automated detection (Meena and Suriya, 2023).

Theme 5: GenAI can improve learning when structured as a scaffold (especially in skills training)

All findings are not negative. Two randomized controlled trials of dental skills education have shown that ChatGPT-based training leads to better performance and perceptions among learners (e.g., self-efficacy, motivation) as compared to control conditions, including but not limited to the fact that AI facilitation provides an improvement in performance and perceptions (Huang et al., 2025). In the same vein, a study involving undergraduate EFL learners (n = 250) through an intervention program that used ChatGPT to support a computer-based writing program focused on critical thinking showed that a critical-thinking-oriented writing program enhanced critical thinking as indicated by the writing dimensions and by the constructs of acceptance (e.g., trust and performance expectancy) that was more accepted (Yin and Dou, 2025). Such outcomes suggest that GenAI can be used to facilitate critical thinking when the learning tasks demand that students evaluate, refine and justify the results instead of submitting AI texts as final work (Huang et al., 2025; Yin & Dou, 2025).

Theme 6: Institutional governance is active but inconsistent across universities

Analysis of policies reveals that institutions are acting but not similarly. A cross-national study of 40 universities (six regions) found various strategies of adoption and disproportional focus on integrity, literacy, and implementation mechanisms (Jin et al., 2024). An independent mixed-method investigation of the guidelines at the best 50 U.S. universities (topic modeling + sentiment + thematic analysis) also found that institutions do focus on flexible, stakeholder-specific policies, with frequently a role of faculty to establish boundaries at the course level (An et al., 2025). The lack of consistency in policy ecosystems may result in confusion of students and discriminatory practices in different departments, which proves the necessity of stricter design of standards and assessment (An et al., 2025; Jin et al., 2024).

Theme 7: Adoption is shaped by stress and productivity pressures integrity norms can suppress use

Research on empirical studies of academic adoption indicates that rational productivity under stressors partially motivates GenAI adoption. A survey-based study (n = 702) revealed time-saving, self-efficacy, and perceived stress to predict adoption, and academic integrity issues influenced them negatively (Bin-Nashwan et al., 2023). This implies that integrity interventions that do not take into consideration workload and performance demands might not be as effective as those that restructure work and offer valid and transparent avenues of AI utilization (Bin-Nashwan et al., 2023).

Discussion (thematic, evidence-led, 2023–Feb 2026)

1) Generative AI changes *what academic performance means*: fluent output no longer equals demonstrated reasoning

One of the main findings in the existing empirical literature is that GenAI increases the likelihood of performance without understanding, i.e., students can write coherent work without necessarily having to go through the cognitions (analysis, synthesis, evaluation) that academic work is intended to test. The best example of actual-life evidence is a blind "Turing test" case study which had penetrated an actual university exam system and inserted fully AI-written submissions in several modules; 94% were not detected by markers, and AI scripts tended to score higher than the genuine student work. This is not a lab simulation, but the direct trial of the operations of institutional assessment (Scarfe et al., 2024).

This observation aids to clarify how the issue of integrity has changed over time with respect to the conventional concept of plagiarism into a more broad-based problem of assessment validity. The credential signal is undermined in case of an inability of routine coursework to reliably differentiate between reasoning of a student and a tool prose (Scarfe et al., 2024).

2) Cognitive offloading can reduce durable learning unless AI use is structured

Critical thinking crisis is not merely ethical but it is also cognitive. Long-term retention was tested in a randomized controlled trial (RCT) testing study where students who used ChatGPT as a study aid scored much worse on a surprise test 45 days later than students who used traditional study methods (Barcaui, 2025).

This is consistent with cognitive offloading: in cases where AI helps to decrease effortful retrieval, elaboration, and self-explanation, students will feel productive but they will develop weaker memory representations.

Simultaneously, the empirical research also indicates that the results are also strongly dependent on the interactions of students with AI. An experimental study (pretest-posttest control group), which is carried out as a mixed-method research among undergraduates, demonstrated significant changes in critical, creative, and reflective thinking when students used ChatGPT as a part of learning (Essel et al., 2024).

And, according to survey + performance measure research in Italy, performance in critical thinking was a better predictor of engagement with AI (not merely knowledge about AI), attitudes and trust influenced engagement, and engagement was associated with critical reasoning performance (Suriano et al., 2024). The Interpretation: GenAI does not necessarily decrease critical thinking: instead, it may lead to a bifurcation: Unorganized application thinking offloading, output acceptance, poor retention (Barcaui, 2025). Critique, verification, justification Structured, active, i.e. potential gains in higher-order reasoning (Essel et al., 2024; Suriano et al., 2024).

3) Integrity has shifted from "copying" to "provenance": disclosure norms and compliance are fragile

Conventional systems of integrity were constructed on duplication of human beings or references. GenAI generates text that has the appearance of an original and also lacks authorship, accelerating institutions towards proving the existence of integrity based on provenance: documenting the contribution of AI. However, studies indicate that disclosure systems are weak as practice. When higher education claimed AIs use declarations, a study has indicated that there is significant non-compliance despite the requirement, claiming that the behavioral motivators (norms, perceived risk, feasibility) can influence whether students disclose (Gonsalves, 2024/2025).

Simultaneously, an equivalent study using theory-based survey evidence demonstrates that perceived risk (privacy, consequences of misconduct) of students negatively correlate with frequency/intention of using ChatGPT, and perceived dishonesty may mediate through risk perception instead of being directly correlated with use (Social Psychology of Education study; n=468) (Blahopoulou & Ortiz-Bonnin, 2025).

Interpretation: disclosure cannot be a policy box item, but rather a social convention. In case of uncertain, inconsistent or even perceived punitive policies, students will hide the usage even in cases of assistance which complicates provenance tracking and equitable application.

4) AI detection tools: empirically inconsistent, ethically risky as primary evidence

One of the great practical temptations is to find the solution to integrity by detectors. However, according to several original studies, detectors are unreliable and error-prone. A study about Thinking Skills and Creativity which concentrated on false positives and false negatives also indicated that false positives are the most dire and certain parts of the work (e.g., literature review-style writing) have a higher likelihood of being marked incorrectly (Dalalah and Dalalah, 2023).

In a different original study of the International Journal of Educational Integrity, a variety of detectors were tested on paragraphs generated by ChatGPT 3.5 and GPT-4 and on human controls; detectors were more accurate on GPT-3.5 than GPT-4, and generated false positives/ambiguous classifications on human text (Elkhatat et al., 2023).

Academic integrity implications: When detectors have non-trivial false positives, they are capable of making false accusations and freezing lawful learning assistance (in particular with multilingual writers). The facts show that it is a model of integrity in which the outputs of detection are at most triage signals- never exclusive evidence (Dalalah and Dalalah, 2023; Elkhatat et al., 2023).

5) Evidence that AI can *support* critical thinking when aligned with pedagogy (interventions and RCTs)

The most notable opposing point to crisis stories is that a number of original studies show benefits with guided designs.

Skills education RCT (dental): This randomized controlled trial involved a combination of ChatGPT and instructional videos and showed better skills and variables of learners (e.g., motivation/self-efficacy) than control and included cognitive-load measures (Huang et al., 2025).

Critical-thinking-based writing treatment (EFL, n=250): A pretest-posttest case study revealed that a critical thinking writing program based on AI-assistance positively influenced writing proficiency and learner approval variables indicating an improvement in cases where learners are forced to engage in the active accomplishment of writing and inference (Yin and Dou, 2025).

The findings pertinent to the research question are the evidence of a scaffold not substitute principle: when tasks are supposed to make students analyze the output of AI, create counterarguments, provide evidence, and contemplate revisions as well, the latter can act as an on-demand tutor instead of an answer engine.

6) Student perceptions at scale: heavy use for efficiency; skepticism about critical-thinking development

Studies on perception at large scales reflect that many students will use ChatGPT to brainstorm, summarize, and source find and that students tend to consider efficiency. An international poll (23,218 students; 109 countries/territories) revealed that students found ChatGPT helpful to simplify and summarize information, but not so helpful to develop critical thinking, and they generally supported the idea of regulation because of cheating/plagiarism issues (PLOS ONE international study, perceptions early 2024) (Kamalov et al., 2025).

A survey of undergraduates in the U.S. also found a high usage rate in understanding concepts and organizing academic assignments, and it highlights the normalized role of AI in the daily coursework (US student perceptions survey) (Musa et al., 2025).

Interpretation: AI is commonly perceived by students as a productivity tool. To ensure that critical thinking is maintained, institutions should design learning tasks that induce instances of slow thinking, when they are checking, challenging, and defending arguments, and not rewarding fluency.

7) Institutional responses: active governance, but inconsistent and decentralized

Analysis of policy indicates that although universities are issuing guidance at an impressive rate, there is a great deal of variation in what they can use, the extent of disclosure expectations, and the model of enforcement. The global policy study implemented in universities revealed a variety of adoption practices and highlighted the imbalance in the detail of implementation (Jin et al., 2024).

An American-oriented study of university policies also exhibits a trend of shifting the burden to course-based policies to facilitate contextual flexibility but enhancing the lack of conformity between departments (An et al., 2025).

Practical implication: the lack of consistent rules produces ambiguity, which in turn causes intentional breaches and abuse (students who do not know what they can do) and intentional abuse.

8) The motivational ecology: adoption is shaped by stress and productivity pressures, not only ethics

Lastly, according to original adoption research, the issue of integrity is structural as well. A massive literature review of scholars/researchers identified time-saving reasons, self-efficacy, and stress as the predictors of adoption, and the academic integrity issues as the negative predictors of use-a long-standing conflict between productivity motivators and ethical principles (Bin-Nashwan et al., 2023).

Interpretation: The moralistic integrity strategy of not using AI will not always work in environments where a workload is high and where support is scarce. Better strategies are made up of work load conscious design (authentic assessments, developed drafts, oral defenses) as well as explicit ethical literacy and realistic allowed-use trajectories.

Conclusion

Generative artificial intelligence (AI) coming to a rapid integration into higher education is not only a change in technology, but a change in the structural production, assessment, and authentication of knowledge. The evidence explored in this research paper indicates that academic integrity and critical thinking present both threats and opportunities in the generative AI. Empirical evidence indicates that the usage of AI (without structure) may result in cognitive offloading, a decrease in long-term knowledge memory, and a decrease in higher-order reasoning. Assessment studies in the real world also indicate that AI-generated work is capable of passing the institutional evaluation systems, which reveals a weakness in the conventional product-based assessment framework. The results of these studies affirm that the existing educational integrity model, which is, to a large extent, built upon plagiarism detection engine and content originality, is not adequate in the age of generative AI.

Nonetheless, it is also evident in the literature that generative AI is not inevitably harmful to the critical thought process. AI can be used as a scaffold and not as a substitute of reasoning when incorporated into systematic pedagogical constructions that must be assessed, justified, reflected upon and revised. According to experimental and intervention studies, guided AI integration has the potential to improve learner interaction, facilitate the acquisition of skills, and, even, improve some aspects of critical thinking. It is not what the presence of AI is, but the circumstances in which it is applied.

It is institutional adaptation, therefore, that is the wider problem. Technologies of detection are not sufficient to protect integrity because of the reliability issues and the ethical aspects of the false positive. Rather, sustainable resolutions need evaluation redesign, express expressive disclosure norms, AI literacy education and consistent governance frameworks. The concept of academic integrity should be developed as not being limited to the idea of ownership of the text but being more concerned with the issues of transparency, responsibility, and provable reasoning practices.

Finally, misalignment between newer tools and older academic frameworks is what leads to the crisis of critical thinking in the era of generative AI, and the technology is not the problem. University education needs the repositioning of human reasoning, ethical responsibility, and intellectual rigor in the centre of AI integration. Provided that institutions manage to reshape their assessment and policy frameworks to focus on authentic cognitive interactions, generative AI potentially will not be a threat to academic integrity, but a driver of twenty-first-century academic integrity redefinition and reinforcement.

Recommendations

1. Redesign assessments to prioritize critical thinking processes (e.g., drafts, reflections, and oral defenses) rather than relying solely on final written outputs.
2. Establish clear institutional policies requiring transparent disclosure of generative AI use in academic work.
3. Integrate AI literacy education into curricula to ensure students critically evaluate and responsibly use AI tools.

References

1. An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*. https://doi.org/10.1186/s41239-025-00507-3
2. Barcaui, A. (2025). *ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. Social Sciences & Humanities Open, 12,* 102287. https://doi.org/10.1016/j.ssaho.2025.102287
3. Bin-Nashwan, S. A., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society, 75,* 102370. https://doi.org/10.1016/j.techsoc.2023.102370

4. Blahopoulou, A., & Ortiz-Bonnin, S. (2025). Chat or cheat? Academic dishonesty, risk perceptions, and ChatGPT usage in higher education students. *Social Psychology of Education.*

5. Coates, H., Croucher, G., & Calderon, A. (2025). Governing academic integrity: Ensuring the authenticity of higher thinking in the era of generative artificial intelligence. *Journal of Academic Ethics, 23,* 2015–2028. https://doi.org/10.1007/s10805-025-09639-7

6. Dalalah, D., & Dalalah, O. M. A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *Thinking Skills and Creativity, 50,* 101401.

7. Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity, 19,* Article 17.

8. Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2023). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence, 4,* 100198. https://doi.org/10.1016/j.caeai.2023.100198

9. Gonsalves, C. (2025). Addressing student non-compliance in AI use declarations: Implications for academic integrity and assessment in higher education. *Assessment & Evaluation in Higher Education, 50*(4), 592–606. https://doi.org/10.1080/02602938.2024.2415654

10. Hsu, H.-Y., Hakouz, A., & Fotouhi, G. (2025). Towards responsible generative AI in academia: A synthesis of AI policies on academic writing in the field of educational research. *AI and Ethics, 5,* 5467–5484. https://doi.org/10.1007/s43681-025-00794-6

11. Huang, S., Wen, C., Bai, X., Li, S., Wang, S., Wang, X., & Yang, D. (2025). Exploring the application capability of ChatGPT as an instructor in skills education for dental medical students: Randomized controlled trial. *Journal of Medical Internet Research, 27,* e68538. https://doi.org/10.2196/68538

12. Hussain, S. (2023). A study on the quality of secondary education in government and Punjab education foundation institutions about sustainable development goal 4 (SDG-4) 2025. *International Research Journal of Education and Innovation, 4*(3), 41-55.

13. Hussain, S. (2024). An Experimental Study on the Impact of Digital Textbooks on the Academic Achievement of Elementary School Students. *International Research Journal of Education and Innovation, 5*(1), 16-27.

14. Hussain, S., & Abbas, Q. (2023). Examine the Secondary Level Administration Quality of Public and Foundation Funded Schools by the Punjab Education Foundation. *International Research Journal of Education and Innovation, 4*(4), 45-59.

15. Hussain, S., & Khoso, A. A. (2021). Examining the Relationship Between Having a Dedicated Study Space at Home and Secondary Students' Academic Achievement. *International Research Journal of Education and Innovation, 2*(3), 337-345.

16. Hussain, S., & Khoso, A. A. (2021). Examining the Relationship Between Access to Home Amenities and Students' Academic Achievement at the Secondary Level. *International Research Journal of Education and Innovation, 2*(3), 325-336.

17. Hussain, S., & Khoso, A. A. (2022). Examining the Relationship Between Homeownership Status and Secondary Students' Academic Achievement. *International Research Journal of Education and Innovation, 3*(1), 398-407.

18. Hussain, S., & Khoso, A. A. (2022). Investigating the Relationship Between Parents' Education and Students' Academic Achievement at the Secondary Level. *International Research Journal of Management and Social Sciences, 3*(1), 352-363.

19. Hussain, S., Khan, Z., & Khan, R. M. A. (2024). Building Students Creative Thinking Ability Through STEM Integrated Curriculum: An Experiment on Elementary School Students. *International Research Journal of Management and Social Sciences, 5*(1), 12-24.

20. Hussain, T., Khan, H. M. R., Saqib, M., & Qahar, A. (2024). The Most Influential Determinants of School Dropout: A Case of Special Education Complex at Skardu. *International Research Journal of Management and Social Sciences, 5*(2), 81-87.

21. Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2024). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8,* 100348. https://doi.org/10.1016/j.caeai.2024.100348

22. Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE.*

23. Meena, S. M., & Suriya, S. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *Thinking Skills and Creativity, 50*, 101401. https://doi.org/10.1016/j.tsc.2023.101401

24. Musa, S., [et al.]. (2025). The use of ChatGPT in academia: Perspectives of higher education students. *Cogent Education.*

25. Oh, S. H., & Sanfilippo, M. R. (2025). Responsible AI in academia: Policies and guidelines in US universities. *Information and Learning Sciences, 126*(9–10), 561–587. https://doi.org/10.1108/ILS-03-2025-0042

26. Perveen, F., & Hussain, S. (2023). Enhancing Teaching Effectiveness: The Significance of Subject Matter Proficiency in Alignment with Pakistan's National Professional Standards for Teachers. *International Research Journal of Management and Social Sciences, 4*(4), 624-633.

27. Peterson, S. (2025). Addressing student use of generative AI in schools and universities through academic integrity reporting. *Frontiers in Education, 10*, 1610836. https://doi.org/10.3389/feduc.2025.1610836

28. Sabir, H. U. S. S. A. I. N., Masood, A. H. M. A. D., Sobia, A. L. T. A. F., & Muhammad, F. A. (2024). Evaluation of Effect of Quality Standards on Quality of Education in Public and Punjab Education Foundation Funded Schools at Secondary Level. *JCTE Учредители: Secondary Teacher Education Department, Allama Iqbal Open University, 7*(1).

29. Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. *PLOS ONE, 19*(6), e0305354. https://doi.org/10.1371/journal.pone.0305354

30. Sheeza, S., Khan, H. M. R., Ijaz, A., & Qureshi, U. F. (2024). Research Method in International Comparative Education. *International Research Journal of Management and Social Sciences, 5*(1), 544-574.

31. Smith, S., Tate, M., Freeman, K., Walsh, A., Ballsun-Stanton, B., Hooper, M., & Lane, M. (2025). A university framework for the responsible use of generative AI in research. *Journal of Higher Education Policy and Management*, 1–20. https://doi.org/10.1080/1360080X.2025.2509187

32. Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2024). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction.*

33. UNESCO. (2023). *Guidance for generative AI in education and research.* https://doi.org/10.54675/EWZM9535

34. Yin, X., & Dou, K. (2025). An AI-assisted critical thinking intervention to enhance undergraduate EFL learners' writing proficiency. *Studies in Educational Evaluation, 86*, 101480. https://doi.org/10.1016/j.stueduc.2025.101480

35. Zafar, N., Khan, H. M. R., Ashraf, S. R., & Nawaz, I. (2024). Teachers' Expectations for Effective and Successful Training Programs: Inspiring Growth and Excellence. *International Research Journal of Management and Social Sciences, 5*(2), 895-910.