# Generative AI in University Teaching: An Exploratory Qualitative Study of Five Teachers at One University

**Bin Yu[1*], Liyuan Wang[2], Wenjie Si[3], Yijun Liu[1], Lihang Ma[1]**

**[1] Pingdingshan University, Pingdingshan, Henan, China**

**[2] Pingdingshan Vocational and Technical College, Pingdingshan, Henan, China**

**[3] Henan University of Urban Construction, Pingdingshan, Henan, China**

**DOI: https://doi.org/10.51244/IJRSI.2026.1310000176**

**Received: 20 October 2026; Accepted: 25 October 2026; Published: 12 November 2026**

## ABSTRACT

Generative artificial intelligence (GenAI) can accelerate the preparation of teaching materials, but its educational value depends on teachers' evaluation of outputs and the conditions supporting their work. This exploratory qualitative descriptive study examined teaching experiences, challenges, and perceived institutional support needs among five teachers at one Chinese university. Participants taught computer science, Chinese language and literature, mathematics, physical education, and accounting. Individual semi-structured interviews were conducted in Chinese and lasted 17–21 minutes. Two researchers jointly coded the interview material and discussed themes. Four preliminary themes concerned pedagogical judgement, context-dependent reliability risks, assessment redesign, and mismatches between institutional provision and practical needs. Teachers described converting erroneous outputs into critical learning activities, checking generated content against disciplinary requirements, and seeking evidence of students' reasoning beyond completed assignments. Other accounts showed limited use or withdrawal when verification costs, uncertainty over grading, or professional accountability outweighed perceived benefits. Assessment responses included oral explanations, process documentation, and supervised tasks, but participants also described additional marking work and reduced opportunities for open-ended learning. Institutional accounts emphasised generic training, unclear assessment procedures, and limited recognition of additional work. Priorities varied across individual cases, including verification resources, practical teaching technologies, credible materials, and usable assessment guidance. These findings suggest that responsible integration requires support for teachers' decisions about when GenAI serves a pedagogical purpose, including decisions to restrict its use. Each discipline was represented by one participant, and the interview accounts were not independently corroborated. The study therefore offers contextual, exploratory insights rather than evidence of disciplinary patterns, measured learning gains, or generalisable institutional effects.

**Keywords:** generative artificial intelligence; university teachers; qualitative interviews; professional judgement; institutional support

## INTRODUCTION

Generative artificial intelligence (GenAI) enables university teachers and students to generate explanations, examples, code, feedback, and assessment materials. Its flexibility creates opportunities for instructional preparation and experimentation, while also raising questions about accuracy, learning, and responsibility (Kasneci et al., 2023; Lo, 2023). A convincing response is not necessarily an appropriate teaching resource. Teachers must establish whether generated content meets course objectives, disciplinary standards, and students' needs before using it.

The resulting work extends beyond learning to operate a tool. Teachers may need to verify references, test code, check mathematical reasoning, or adapt activities to physical settings. Student access also complicates the interpretation of completed assignments: fluent writing or functioning code may offer limited evidence

of independent understanding (Luo, 2024; Xia et al., 2024; Yan et al., 2024). These challenges connect instructional decisions with institutional arrangements governing resources, assessment, professional development, and workload.

Existing research already documents educators' concerns and support needs. The present study therefore does not claim that teachers' experiences have been overlooked. Instead, it examines how concrete teaching episodes connect perceived benefits, verification demands, assessment responses, and institutional support within one setting. Particular attention is given to selective use and withdrawal, which can reveal constraints obscured by an emphasis on adoption alone.

Five teachers from different subjects at the same university were interviewed. The study asks: RQ1: How do these teachers experience and integrate GenAI into teaching? RQ2: What challenges arise in their particular instructional contexts? RQ3: What institutional support do they consider necessary? With one participant per subject, the purpose is exploratory case comparison rather than disciplinary representation. The contribution is a contextual account of how professional judgement mediates the relationship between technological possibilities and workable teaching practices.

## LITERATURE REVIEW

### Teaching Applications and Professional Judgement

Reviews describe varied pedagogical applications of GenAI, including planning, resource development, feedback, and learning activities (Qian, 2025; Wang et al., 2025). Work on lesson planning similarly identifies opportunities for generating alternatives and supporting critical thinking (Van Den Berg & Du Plessis, 2023). These possibilities depend on teachers' adaptation of outputs; generating more material does not itself demonstrate better instruction. Zhai's (2025) account of teachers' roles and agency highlights the relationship between knowledge, engagement, and professional support.

Reliability introduces a further layer of judgement. The educational challenges reviewed by Yan et al. (2024) include practical and ethical problems surrounding large language models. In a qualitative study of 20 pre-service science teachers, Hu et al. (2026) examined how participants assessed GenAI claims through interviews and vignette-based tasks. Their findings connected standards of evidence and verification with the stakes of intended use, especially teaching. That study concerns prospective science teachers; it does not establish how experienced teachers across different subjects manage verification within everyday workloads.

### Assessment and Institutional Conditions

Reviews of assessment and institutional policies identify tensions around permissible assistance, originality, and evidence of learning (Luo, 2024; Xia et al., 2024). Detection provides an uncertain basis for resolving these tensions. Liang et al. (2023) demonstrate bias affecting non-native English writing; their findings should not be assumed to quantify the performance of the Chinese-language systems encountered here. Conceptual work on AI-resilient assessment instead foregrounds the evidence required to justify judgements of competence (Nikolić & Basta Nikolić, 2026).

Institutional support also involves more than access. Research addresses faculty development needs (Mah & Groß, 2024) and professional GenAI competence in language teacher education (Moorhouse et al., 2024). The distinction between faster production and additional academic work is explicit in the workload review by Su et al. (2026). UNESCO's guidance places human agency, privacy, and pedagogical validation within a broader framework for educational use (Miao & Holmes, 2023). These perspectives direct attention to whether institutional provision helps teachers make defensible decisions in their own courses.

### Focus of the Present Study

Lee et al. (2024) investigated educators' perspectives through a staff survey and interviews at an Australian university, identifying varied responses and demand for support. Together with the qualitative work of Hu et

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

al. (2026), this provides a more specific starting point than a general claim that teacher experience is under-researched. The present inquiry connects teachers' reported classroom incidents with decisions to continue, limit, or discontinue particular uses at one Chinese university. It explores how verification costs, assessment uncertainty, and perceived institutional recognition interact across five individual accounts. It cannot isolate disciplinary effects, measure effectiveness, or establish a new general model of adoption.

# METHODOLOGY

## Design and Setting

An exploratory qualitative descriptive design was used to examine teachers' accounts of GenAI-assisted teaching. All participants worked at the same university in China. One participant described the institution as a teacher-education university. Participants reported shared accounts, training, AI functions within the teaching platform, and a trial usage document. These are interview accounts, not independently verified descriptions of institutional provision. Their interpretations of policy differed: some emphasised encouragement and broad principles, while others emphasised restrictions and detection. The study therefore examines perceived support rather than policy implementation itself.

## Participants and Data Collection

Purposive sampling recruited five teachers with teaching-related GenAI experience. P01 taught computer science, P02 Chinese language and literature, P03 mathematics, P04 physical education, and P05 accounting. Teaching experience ranged from 6 to 16 years. Individual face-to-face interviews lasted approximately 19, 21, 17, 18, and 20 minutes, respectively, totalling approximately 95 minutes. The supplied interview records state that audio recording occurred with consent and was followed by transcription. Chinese was used throughout the interviews. The common guide covered six areas: tools and teaching uses; a concrete teaching episode; perceived changes in work; difficulties and responses; existing institutional support; and priorities for improvement. Follow-up prompts sought details about errors, time costs, assessment decisions, and the reasons for suggested support. An illustrative prompt asked participants to describe a specific occasion when AI assisted preparation or classroom teaching; another asked which single institutional action they would prioritise. This structure elicited contrasting experiences, including unsuccessful use, rather than benefits alone.

These brief interviews support a focused exploratory account. Subject variation provides contrasting examples but little depth within any discipline. No claim of saturation is made. The dataset is adequate for identifying issues for further inquiry, not for mapping all relevant experiences or estimating how common they are.

## Analysis and Evidential Transparency

The interview material was examined using descriptive thematic analysis. Two researchers jointly coded the material and discussed the themes. The analysis is reported at the level of descriptive patterns across the five accounts, drawing on thematic analysis as a flexible approach to qualitative interpretation (Braun & Clarke, 2006). The four themes concern instructional judgement, reliability, assessment, and institutional support.

Participant-labelled excerpts show the relationship between the accounts and the interpretations. Contrasting cases, including P03's limited use and P05's retreat from substantive applications, remain visible within the thematic account. Joint coding and discussion supported interpretive deliberation; they should not be understood as independent reliability testing. The study does not claim member checking, a formal audit trail, or external triangulation.

## Translation and Ethical Considerations

Selected Chinese excerpts are presented in English with participant identifiers retained. AI assistance was used in drafting translations, language editing, and manuscript restructuring. The translated quotations were

compared with the Chinese wording during preparation of this revision. The eleven excerpts preserve the substantive meaning of the selected passages; no back-translation or participant validation of the English translations is claimed. Translation nevertheless introduces an additional layer of interpretation.

The study received ethical approval from Pingdingshan University. All five participants signed written informed consent forms that included permission to publish anonymised interview quotations. Interviews were audio-recorded with consent. Participant codes are used in reporting the findings, and identifying details are limited to those necessary for interpreting the accounts.

Identifiers P01–P05 replace names. Exact courses, academic rank, and individual service histories are omitted from the findings to reduce identification risk. Unique subject roles and recognisable incidents may nevertheless permit identification within a single institution; pseudonyms alone cannot guarantee anonymity. The findings therefore avoid naming the institution and restrict personal detail to information relevant to interpreting the accounts.

## FINDINGS

Four preliminary themes link the three research questions. Theme 1 addresses teaching experiences (RQ1), Theme 2 contextual challenges (RQ2), Theme 3 experiences of assessment change (RQ1–RQ2), and Theme 4 support needs (RQ3). Accounts are teachers' recollections, including their interpretations of student responses; classroom effects were not independently measured. Subject labels identify individual cases, not representative disciplinary positions.

### From Generating Content to Exercising Pedagogical Judgement

P01 used GenAI to develop a shortest-path lesson around a transport network. The generated explanation incorrectly extended Dijkstra's algorithm to graphs with negative-weight edges. After testing the material, P01 converted its weaknesses into a group task: students examined the explanation and code for errors. P01 described a subsequent change in practice: “I basically no longer take AI output directly into the classroom. I let it produce something first, then review it myself.” This account connects generation with checking and task design, rather than straightforward substitution for preparation.

P02 juxtaposed three AI-generated passages with a student's writing in a lesson on descriptive detail. Students discussed why polished language could appear human and why an awkward but personally situated detail mattered. The activity led to an assignment requiring an irreplaceable detail and an explanation of its source. P02 described a shift in writing instruction: “What they lack is not output, but judgement.” The participant also acknowledged that some students questioned the value of repeated revision when a new text could be generated quickly. The reported innovation therefore involved continuing pedagogical disagreement. The pattern was not universal. P03 stated: “It changed how I prepare, but not how I teach.” AI supplied possible analogies, while classroom explanation and stepwise reasoning remained largely unchanged. P04 similarly described faster production of lesson-plan text, with saved time redirected towards organisation and safety. These accounts distinguish different kinds of change: generating alternatives, redesigning learning activities, and assisting administrative preparation did not amount to the same teaching transformation.

P02 also questioned an asymmetry in disclosure: students were required to declare assistance and retain drafts, while the teacher had not disclosed the extent of personal preparation use. The participant regarded the justification that teachers used AI to improve teaching while students used it to replace learning as inadequate. This self-criticism complicates an account in which teachers simply supervise student misuse; professional judgement also involved examining their own expectations and practices.

### Reliability Judged Against the Consequences of Error

For P03, attractive formatting concealed mathematical and curricular problems. Of twelve generated calculus variations, only four were used after checking conditions, solutions, diagrams, and syllabus

alignment. Verification took two evenings. P03 concluded: “It is not that it cannot be used; the return on the effort is not worthwhile.” The participant remained receptive to initial ideas but avoided relying on AI to guarantee correctness, particularly during live teaching. P01 and P02 also reported fabricated references, which prompted direct checks against retrievable sources and original texts.

P04 described two risks in practical teaching: a mnemonic reversing a movement sequence and a training plan omitting recovery days. Neither account establishes that an injury occurred. The concern was that unreviewed guidance could become embodied practice: “Once a movement is learned incorrectly, it becomes ingrained in the body.”

Yet P04 also reported a positive case: AI suggested alternatives for a student unable to participate in ordinary running and jumping activities. The teacher checked sources, consulted a campus doctor, and adapted the suggestions. The account supports supervised adaptation, not the safety of AI-generated exercise prescriptions.

P05's internally inconsistent financial statements disrupted a classroom case. The error undermined confidence because the lesson itself concerned relationships among statements. The participant subsequently limited teaching-related use largely to translation and document reading. P05 explained the professional concern: “Every choice has someone behind it who must take responsibility.” This was a personal response to a reported incident and assessment difficulties, not evidence that accounting teachers generally withdraw from GenAI.

Across these cases, reliability meant more than a correct-looking answer: it involved valid reasoning, faithful source representation, workable physical activities, or coherent financial evidence. The consequences of error informed decisions about acceptable uses. However, the study cannot determine whether subject, experience, personal confidence, or the particular incident best explains these differences.

**Seeking Evidence of Learning Beyond the Submitted Product**

P01 described students submitting sophisticated code that they could not explain. Rather than automatically treating this as misconduct under unclear rules, the teacher added short oral explanations and requested drafts or AI interaction records. The revised criterion was expressed as: “Using AI is not the problem; being unable to explain it is the problem.” This sought evidence of understanding while allowing assistance, though the reported improvement was not independently verified.

P02 encountered a student's writing flagged by a detection system. Earlier drafts and notes supported the student's authorship, and the teacher graded it as original. The episode exposed additional work needed to challenge a suspicious score: “I have to spend time proving that a student did not cheat.” P02 introduced some handwritten classroom tasks but recognised that this reduced opportunities for longer assignments requiring research. The incident illustrates a reported conflict between procedural reassurance and the range of learning activities available; it does not establish the detector's overall error rate.

P03 responded to perceived differences between homework and supervised performance by requiring key solution steps and short classroom quizzes. P05 described an appeal over similar financial analysis reports and difficulty articulating defensible grading distinctions. Open-ended coursework was replaced by more controlled assessment, alongside an oral component. Together, these cases suggest that uncertainty concerned both learning evidence and the justification of marks. The responses offered closer access to reasoning, but increased workload or narrowed tasks. They should not be interpreted as proof that supervised formats are always preferable.

**Support that Fits Teachers' work**

Participants consistently described basic institutional provision but differed in their priorities. P01 sought peer-led teaching cases, reliable infrastructure, and sufficient access for course-scale use. P02 wanted credible text resources, copyright guidance, and fair assessment procedures. P03 prioritised mathematical

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

verification: “What I want is ‘how to make it calculate correctly’; what they explain is ‘how to make it sound convincing’.” This contrast concerned the participant's experience of training, not an objective evaluation of every institutional session.

P04 sought movement analysis, video annotation, and other tools suitable for practical classes. Some of these are not generative systems, which highlights a mismatch between a GenAI promotion agenda and perceived teaching needs. The participant also reported difficulty attending training alongside practical teaching commitments. P05 prioritised usable evidence and grading criteria: “First give us a basis for making judgements, rather than tools first.” Across cases, broad policy language did not resolve course-level uncertainty about disclosure, acceptable assistance, or responsibility.

Workload recognition linked these requests. Participants described time spent testing outputs, redesigning assignments, learning changing tools, or attending training, with limited perceived recognition in evaluation. P05 summarised withdrawal: “It is not that I figured something out; I worked out the costs.” Several participants opposed judging teaching quality by usage frequency and wanted permission to decline unsuitable applications. Thus, support meant the capacity to make and defend an appropriate choice, including restricted use, rather than encouragement alone.

## DISCUSSION

### Integration as a Conditional Professional Decision

The findings connect opportunities for generation with teachers' responsibilities for instructional selection. P01 and P02 described using generated material to cultivate critical evaluation; P03 retained AI largely as a source of preparatory ideas; P05 restricted substantive use. This variation qualifies broad accounts of changing teacher roles (Van Den Berg & Du Plessis, 2023; Zhai, 2025). Within this small setting, engagement followed different trajectories, and greater use did not necessarily indicate a more thoughtful response. The data support attention to decisions about particular tasks rather than a single adoption category for each teacher.

The study also extends the local question beyond whether outputs are trustworthy. In the accounts, teachers considered what checking would require, what an error could do, and whether they could justify a decision afterwards. This interpretation connects with the attention to evidence and stakes in the work of Hu et al. (2026), while adding experienced teachers' reported constraints in actual courses. It remains an interpretation of interview accounts, not a causal explanation or a validated model of professional judgement.

### Verification and Assessment Redistribute Work

Reports of faster drafting coexisted with substantial checking and redesign. This supports examining the whole task when considering efficiency, consistent with the workload question raised by Su et al. (2026). P03's verification burden and P05's calculation of professional incentives suggest why reduced use may be a reasoned response. Institutions evaluating benefits should examine the time required to obtain usable materials and defensible assessments, not only initial generation speed.

Assessment changes likewise involved competing educational aims. Oral explanations and drafts can make reasoning more visible, aligning with assessment literature's emphasis on evidence of learning (Nikolić & Basta Nikolić, 2026; Xia et al., 2024). Yet controlled formats can restrict research-based work, and process evidence takes time to review. P02's account also warns against treating a detection score as sufficient grounds for a decision. It illustrates a local concern about fairness without extending English-language detector research to an untested Chinese context (Liang et al., 2023).

The practical implication is to align tasks, disclosure requirements, and marking criteria before disputes arise. A course may permit certain assistance while requiring students to explain their reasoning and acknowledge its use. Such an approach needs workload support and clear appeal procedures. The interviews do not demonstrate that any single assessment design prevents misuse or improves learning.

**Institutional Support and the Legitimacy of Limited Use**

The shared setting helps relate contrasting needs to broadly similar reported provision, but does not control for personal or departmental differences. Generic training appeared insufficient for participants seeking source verification, mathematical checking, practical feedback, or defensible grading. Faculty development research (Mah & Groß, 2024; Moorhouse et al., 2024) and UNESCO's human-centred guidance (Miao & Holmes, 2023) provide a rationale for connecting resources with pedagogical purposes rather than promoting tools in isolation.

Tentative priorities include peer workshops based on actual courses, subject-relevant checking resources, accessible training schedules, and clear examples of permissible assistance and assessment evidence. Recognition should cover the time needed to evaluate outputs and redesign tasks. Resources beyond GenAI may be appropriate where teachers need observation or verification rather than further generation. These are proposals arising from perceived needs; their effectiveness and cost have not been evaluated here.

Supporting selective non-use is equally relevant. P03 and P05 did not reject every application; they drew boundaries around particular tasks. Institutions could assess whether those boundaries support learning and accountability rather than treating frequent use as a proxy for innovation. Further consultation with students and policy staff would be needed to turn these interview-based recommendations into institutional procedures.

## LIMITATIONS AND FUTURE RESEARCH

Five short interviews at one university offer limited contextual depth. One teacher per subject prevents separation of disciplinary patterns from individual differences, career experience, or salient incidents. Without longitudinal follow-up, participants' retrospective accounts cannot establish how practices changed over time. Self-reported outcomes, institutional policies, and student experiences were not independently corroborated. Joint coding supports discussion but cannot eliminate interpretive influence. The researchers' relationships with participants and the influence of their own GenAI experiences are not characterised in the available methodological account, which further limits reflexive transparency.

Translation may alter nuance, and residual identifiability persists despite participant codes. The study claims neither saturation nor statistical generalisability. Future research could include repeated interviews, several teachers within each subject, institutions with different support arrangements, observations, and student perspectives. Such evidence could test whether the proposed supports reduce work, protect fair assessment, and enable appropriate teaching decisions.

## CONCLUSION

This exploratory study of five teachers at one university suggests that GenAI integration depends on what teachers can verify, adapt, and defend in particular teaching situations. Opportunities for preparing resources and designing critical learning activities coexisted with concerns about reasoning, authenticity, safety, assessment, and accountability. Selective use and withdrawal were part of the reported experience, alongside experimentation.

The contribution lies in connecting concrete incidents with teachers' perceived support needs. Relevant verification resources, transparent assessment guidance, accessible professional learning, and recognition of additional work may support responsible choices. These implications remain tentative because the data are brief, self-reported, and limited to one case per subject. The study offers questions for institutional deliberation and further investigation rather than evidence for universal disciplinary differences or an obligation to adopt GenAI.

**Appendix Interview Guide**

The following English translations reproduce the six core question areas in the Chinese interview records. Follow-up questions varied with each participant's account. The guide invited both useful and unsuccessful experiences; it was not a scale or a standardised measure.

- Which generative AI tools do you currently use, and for which teaching-related tasks?
- Can you describe a specific experience of using AI to support lesson preparation or classroom teaching? Please describe what happened in detail.
- What practical changes has AI brought to your teaching work?
- What difficulties or problems have you encountered while using it? Please give specific examples.
- What support does your university currently provide for AI-assisted teaching, and does it meet your needs?
- If the university wishes to support further use, what resources or measures are most needed, and why?

Prompts asked participants how they discovered errors, what checking was required, how they responded to assessment difficulties, and which support measure they would prioritise. Interviews ended with an invitation to add anything important that had not been covered.

**Ethical Approval**

The study received ethical approval from Pingdingshan University. All five participants provided signed written informed consent, including permission to publish anonymised interview quotations.

**Conflict of Interest**

The authors declare no conflicts of interest.

**Data Availability**

The underlying interview records are not publicly available because the small, single-institution sample and distinctive teaching incidents create risks of re-identification. Supporting translated excerpts are included in the article. No unrestricted data-access commitment is made; any sharing of the underlying material would require appropriate participant consent and confidentiality safeguards.

## REFERENCES

1. Hu, H., Yu, S., & Xu, L. (2026). Evidence, trust, and objectivity with generative AI: a qualitative interview study of pre-service science teachers' truth-assessment practices. Frontiers in Psychology, 17, 1781772. https://doi.org/10.3389/fpsyg.2026.1781772
2. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (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
3. Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. Computers and Education: Artificial Intelligence, 6, 100221. https://doi.org/10.1016/j.caeai.2024.100221
4. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7), 100779. https://doi.org/10.1016/j.patter.2023.100779
5. 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

6. 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(5), 651–664. https://doi.org/10.1080/02602938.2024.2309963
7. Mah, D.-K., & Groß, N. (2024). Artificial intelligence in higher education: exploring faculty use, self-efficacy, distinct profiles, and professional development needs. International Journal of Educational Technology in Higher Education, 21, Article 58. https://doi.org/10.1186/s41239-024-00490-1
8. Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. System, 125, 103399. https://doi.org/10.1016/j.system.2024.103399
9. Nikolić, D., & Basta Nikolić, M. (2026). Designing AI-resilient assessment in higher education: a four-pillar conceptual framework. Frontiers in Artificial Intelligence, 9, 1841682. https://doi.org/10.3389/frai.2026.1841682
10. Qian, Y. (2025). Pedagogical Applications of Generative AI in Higher Education: A Systematic Review of the Field. TechTrends, 69(5), 1105–1120. https://doi.org/10.1007/s11528-025-01100-1
11. Su, F., Beh, W. F., & Cheah, K. S. (2026). Efficiency or intensification? A systematic review of generative AI and academic workload in higher education. Acta Psychologica, 268, 107353. https://doi.org/10.1016/j.actpsy.2026.107353
12. Van Den Berg, G., & Du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. Education Sciences, 13(10), 998. https://doi.org/10.3390/educsci13100998
13. Wang, P., Jing, Y., & Shen, S. (2025). 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, 65, 100996. https://doi.org/10.1016/j.iheduc.2025.100996
14. Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. International Journal of Educational Technology in Higher Education, 21, Article 40. https://doi.org/10.1186/s41239-024-00468-z
15. Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. British Journal of Educational Technology, 55(1), 90–112. https://doi.org/10.1111/bjet.13370
16. Zhai, X. (2025). Transforming Teachers' Roles and Agencies in the Era of Generative AI: Perceptions, Acceptance, Knowledge, and Practices. Journal of Science Education and Technology, 34(6), 1323–1333. https://doi.org/10.1007/s10956-024-10174-0
17. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
18. Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535