

How Generative AI Is Reshaping Assessment in K12 Informatics Education?

Nguyen Thi Diep Hong¹, Nguyen Thi Dung²

^{1,2} Faculty of Electrical, Electronic and Information Technology, College of Industry, Trade and Tourism

ABSTRACT: Generative artificial intelligence (GenAI) is challenging established assessment practices by enabling students to rapidly generate code, written responses, and other assessable outputs. This systematic review examines how GenAI is reshaping assessment in K-12 informatics education, with particular attention to changes in assessment design, emerging applications, competencies, and risks. The review synthesized 50 studies selected through a multi-stage search, relevance screening, deduplication, and eligibility assessment. Because direct evidence from K-12 informatics remains limited, the corpus also included closely related studies from computing education, K-12 education, STEM, and assessment research. The findings indicate a shift from product-oriented assessment toward approaches that foreground learning processes, higher-order thinking, verification, and students' ability to critically evaluate and justify AI-assisted work. GenAI also shows potential for formative feedback, adaptive assessment, item generation, and assessment design, particularly when implemented with teacher oversight and domain-specific customization. At the same time, its use raises concerns about assessment validity, over-reliance, originality, equity, and whether improved performance reflects students' independent competence. AI literacy, including output verification, critical evaluation, responsible use, and self-reflection, is consequently emerging as an important assessment outcome. Overall, GenAI appears to be driving a substantive reconfiguration of assessment in informatics education, although direct empirical evidence from K-12 informatics classrooms remains limited.

KEYWORDS: generative AI; K-12 education; informatics education; assessment; computer science education; AI literacy; formative assessment

1. INTRODUCTION

Generative AI is reshaping assessment in K-12 informatics education by pushing teachers away from static written products and toward process-focused, higher-order, AI-aware assessment. School-focused evidence is still thinner than higher-education evidence, but across reviews, school studies, and assessment-framework papers, the direction is consistent: coding and essay-like tasks are especially exposed because GenAI can produce plausible answers quickly, while formative feedback, adaptive assessment, and assessment design support are the most promising uses (Chiu, 2024; Kizilcec et al., 2024; Yan et al., 2024).

In K-12 settings specifically, the literature emphasizes two simultaneous shifts. First, AI is being used to generate items, align content to standards, critique assessment quality, and support personalized feedback, which can reduce teacher workload and improve consistency when systems are customized and human-supervised (Bolender et al., 2024; Webb & Galamba, 2026). Second, GenAI destabilizes the validity of conventional take-home assignments, especially written and coding tasks, so assessment increasingly needs to evaluate critical thinking, verification, self-regulation, and ethical AI use rather than final answers alone (Hao et al., 2024; Yan et al., 2024; Zhou et al., 2025).

2. METHODS

The search covered a large multidisciplinary scholarly database containing records from Semantic Scholar, PubMed, and other academic sources. The initial search retrieved approximately 2.9 million records. To improve the specificity and coverage of the search, the initial query was supplemented with targeted search-term reformulations and citation-based searching focusing on combinations of terms related to generative AI, assessment, K-12 education, informatics, computer science education, programming education, and AI literacy.

Following relevance filtering, 240 potentially relevant records were retained for further screening. After duplicate records and studies that did not address the review focus were excluded, 119 publications were considered potentially eligible for detailed assessment. Studies were evaluated according to their relevance to the review question, with priority given to empirical studies and reviews addressing GenAI-supported or GenAI-affected assessment practices in K-12 informatics and closely related educational contexts. The final synthesis included 50 studies that provided the most directly relevant evidence for the review objectives.

How Generative AI Is Reshaping Assessment in K12 Informatics Education?

Studies were considered relevant when they addressed one or more aspects of assessment affected by GenAI, including assessment design, coding or programming assessment, formative feedback, adaptive assessment, assessment validity and authenticity, academic integrity, or competencies required for evaluating and using AI-generated outputs. Particular emphasis was placed on studies involving K–12 informatics, computer science, and programming education.

Because empirical research specifically examining GenAI and assessment in K–12 informatics remains limited, the review also incorporated closely related evidence from K–12 education more broadly, computing and programming education, STEM education, teacher-perception research, and assessment-focused studies in higher education. Evidence from these adjacent contexts was included only when it provided direct conceptual or empirical implications for informatics assessment, particularly in relation to coding assessment, formative feedback, authenticity, validity, higher-order thinking, and AI literacy.

The included studies were synthesized thematically to identify recurring changes in assessment practices and emerging areas of concern. The analysis focused on three overarching dimensions: (1) shifts in assessment design, (2) emerging productive uses of GenAI in assessment, and (3) new competencies and risks associated with AI-mediated assessment. Greater interpretive weight was given to evidence directly situated in K–12 informatics contexts, while findings transferred from adjacent educational settings were treated more cautiously.

3. RESULTS

3.1. Assessment Design Shifts

Across school and higher-education studies, traditional assessment is increasingly seen as misaligned with the GenAI environment, especially when tasks can be completed by chatbots with little visible evidence of student reasoning (Ahangama, 2026; Kizilcec et al., 2024; Weng et al., 2024). Coding and essay assessments appear particularly affected, which matters directly for K-12 informatics because code generation is now a routine GenAI capability (Kizilcec et al., 2024).

Table 1. Emerging Shifts in K-12 Informatics Assessment in the Generative AI Era

Dimension	Vulnerable Tasks	More Robust Alternatives	Citations
Product emphasis	Take-home essays, code outputs	Process evidence, staged tasks	(Kofinas et al., 2025)
Cognitive demand	Recall, routine execution	Analysis, evaluation, creation	(Nguyen Thanh et al., 2023)
Interaction mode	Asynchronous written work	In-person, dialogic, performative	(Kofinas et al., 2025)
AI policy	Unclear prohibition	Transparent permitted use levels	(Perkins et al., 2025)
Learning goal	Final answer correctness	Critical verification and editing	(Hao et al., 2024)

A consistent implication is that K-12 informatics assessment is shifting from “Can the student produce code or text?” to “Can the student reason with, critique, refine, and justify work done with AI?” (Oliveira et al., 2025; Pallant et al., 2026).

3.2. Emerging Productive Uses

The strongest practical opportunities are in formative and adaptive assessment rather than high-stakes autonomous scoring. Reviews and empirical studies describe GenAI as useful for personalized feedback, adaptive questioning, scenario generation, portfolio support, and identifying misconceptions, especially when teachers configure prompts, rubrics, and oversight (Lee et al., 2025; Yan et al., 2024).

K-12-specific work shows similar promise on the assessment-development side. Customized systems improved item generation efficiency, psychometric quality, and standards alignment, while GenAI critique tools produced NGSS-related annotations that overlapped with teacher judgments and supported revision of K-12 assessments (Bolender et al., 2024; Nguyen Thanh et al., 2023).

These gains depend on customization and human-in-the-loop design. Off-the-shelf prompting is not expected to reproduce the benefits seen in specialized educational systems built with psychometricians and domain experts (Bolender et al., 2024; Zapata-Rivera et al., 2024).

3.3. New Competencies and Risks

The K-12 literature increasingly treats AI literacy as an assessable outcome, not just a background skill. Proposed constructs include prompt engineering, factual verification, bias recognition, output evaluation, editing, self-reflection, and ethical use (Hao et al., 2024; Zhou et al., 2025).

At the same time, the evidence warns that GenAI can create a performance illusion: students may perform better with support while developing weaker independent capability, and younger learners remain in a stage where assessing human cognition and metacognition is especially important (Yan et al., 2024). K-12 studies also raise concerns about deception, over-reliance, originality loss, and equity between students with different tool access or executive-function needs (Foung et al., 2024; Lin & Tan, 2025; Zapata-Rivera et al., 2024).

How Generative AI Is Reshaping Assessment in K12 Informatics Education?

4. DISCUSSION

The overall evidence supports a real assessment shift, but its strength varies sharply by claim. The best-supported claims are broad and structural: GenAI disrupts traditional written and coding assessments, encourages more process-centered and higher-order designs, and expands opportunities for formative feedback and personalization (Xia et al., 2024; Yan et al., 2024). Those conclusions are repeated across systematic reviews, scoping reviews, teacher surveys, and conceptual assessment papers, including several that explicitly discuss K-12 or use K-12 cases (Chiu, 2024; Yan et al., 2024).

The weaker part of the evidence is direct proof from K-12 informatics classrooms. Much of the corpus comes from higher education, general school settings, science, or broader AI-in-assessment research rather than K-12 computer science specifically (Chiu, 2024; González-Calatayud et al., 2021; Santos & Junior, 2024). Even when empirical effects look promising, they usually concern adjacent contexts such as Python programming in undergraduates, online computing courses, science-assessment critique, or assessment-item development pipelines rather than day-to-day K-12 informatics classrooms (Lee et al., 2025; Nguyen Thanh et al., 2023).

Still, the direction for K-12 informatics is fairly clear. Because multiple correct solutions are normal in programming, personalized feedback and flexible evaluation are especially attractive, but so are concerns about validity, fairness, and verifying whose thinking is being measured (Arslan et al., 2024; Kaldaras et al., 2024; Qian et al., 2026). That makes human oversight, transparent AI-use policies, and assessment literacy central rather than optional (Khlaif et al., 2024; Moorhouse et al., 2023; Webb & Galamba, 2026).

5. CONCLUSION

Generative AI is reshaping assessment in K-12 informatics education most clearly by making traditional output-based tasks less trustworthy and by expanding the appeal of AI-supported formative, adaptive, and design-oriented assessment. The evidence is strongest for the broad shift in assessment philosophy and weaker for specific K-12 informatics implementations, because direct classroom studies remain sparse.

REFERENCES

1. Ahangama, N. (2026). Designing assessments in the generative AI era: A tailored assessment framework for ICT tertiary education. *International Journal of Educational Technology in Higher Education*, 23(1), 9. <https://doi.org/10.1186/s41239-026-00582-0>
2. Arslan, B., Lehman, B., Tenison, C., Sparks, J. R., López, A. A., Gu, L., & Zapata-Rivera, D. (2024). Opportunities and challenges of using generative AI to personalize educational assessment. *Frontiers in Artificial Intelligence*, 7, 1460651. <https://doi.org/10.3389/frai.2024.1460651>
3. Bolender, B., Vispoel, S., Converse, G., Koprowicz, N., Song, D., & Osaro, S. (2024). Generative AI in K12: Analytics From Early Adoption. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*. <https://doi.org/10.21031/epod.1539710>
4. Chiu, T. K. F. (2024). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10), 6187–6203. <https://doi.org/10.1080/10494820.2023.2253861>
5. Fount, D., Lin, L., & Chen, J. (2024). Reinventing assessments with ChatGPT and other online tools: Opportunities for GenAI-empowered assessment practices. *Computers and Education: Artificial Intelligence*, 6, 100250. <https://doi.org/10.1016/j.caeai.2024.100250>
6. González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial Intelligence for Student Assessment: A Systematic Review. *Applied Sciences*, 11(12), 5467. <https://doi.org/10.3390/app11125467>
7. Hao, J., Von Davier, A. A., Yaneva, V., Lottridge, S., Von Davier, M., & Harris, D. J. (2024). Transforming Assessment: The Impacts and Implications of Large Language Models and Generative AI. *Educational Measurement: Issues and Practice*, 43(2), 16–29. <https://doi.org/10.1111/emip.12602>
8. Kaldaras, L., Akaze, H. O., & Reckase, M. D. (2024). Developing valid assessments in the era of generative artificial intelligence. *Frontiers in Education*, 9, 1399377. <https://doi.org/10.3389/feduc.2024.1399377>
9. Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Ayyoub, A., Hattab, M. K., & Shadid, F. (2024). University Teachers' Views on the Adoption and Integration of Generative AI Tools for Student Assessment in Higher Education. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>
10. Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Radulescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence*, 7, 100269. <https://doi.org/10.1016/j.caeai.2024.100269>

How Generative AI Is Reshaping Assessment in K12 Informatics Education?

11. Kofinas, A. K., Tsay, C. H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, 56(6), 2522–2549. <https://doi.org/10.1111/bjet.13585>
12. Lee, J., Hung, J.-T., Soylu, M. Y., Popescu, D., Cui, C. Z., Grigoryan, G., Joyner, D. A., & Harmon, S. W. (2025). *Socratic Mind: Impact of a Novel GenAI-Powered Assessment Tool on Student Learning and Higher-Order Thinking* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2509.16262>
13. Lin, X., & Tan, H. (2025). A Systematic Review of Generative AI in K–12: Mapping Goals, Activities, Roles, and Outcomes via the 3P Model. *Systems*, 13(10), 840. <https://doi.org/10.3390/systems13100840>
14. Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.cao.2023.100151>
15. Nguyen Thanh, B., Vo, D. T. H., Nguyen Nhat, M., Pham, T. T. T., Thai Trung, H., & Ha Xuan, S. (2023). Race with the machines: Assessing the capability of generative AI in solving authentic assessments. *Australasian Journal of Educational Technology*, 39(5), 59–81. <https://doi.org/10.14742/ajet.8902>
16. Oliveira, M., Zednik, C., Bombaerts, G., Sadowski, B., & Conijn, R. (2025). Assessing students' DRIVE: A framework to evaluate learning through interactions with generative AI. *Computers and Education: Artificial Intelligence*, 9, 100497. <https://doi.org/10.1016/j.caeai.2025.100497>
17. Pallant, J. L., Blijlevens, J., Campbell, A., & Jopp, R. (2026). Mastering knowledge: The impact of generative AI on student learning outcomes. *Studies in Higher Education*, 51(4), 714–735. <https://doi.org/10.1080/03075079.2025.2487570>
18. Perkins, M., Jasper, R., & Furze, L. (2025). Reimagining the Artificial Intelligence Assessment Scale: A refined framework for educational assessment. *Journal of University Teaching and Learning Practice*, 22(7). <https://doi.org/10.53761/rrm4y757>
19. Qian, M., Ji, H., & Li, L. (2026). Empirical validation of a generative AI framework for personalized education assessment. *Scientific Reports*, 16(1), 11538. <https://doi.org/10.1038/s41598-026-42169-9>
20. Santos, S., & S. Junior, G. (2024). Opportunities and Challenges of AI to Support Student Assessment in Computing Education: A Systematic Literature Review. *Proceedings of the 16th International Conference on Computer Supported Education*, 15–26. <https://doi.org/10.5220/0012552500003693>
21. Webb, M., & Galamba, A. (2026). Transforming pedagogy with GENAI -supported formative assessment: Challenges for teacher education. *British Journal of Educational Technology*, 57(3), 690–706. <https://doi.org/10.1111/bjet.70051>
22. Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.9540>
23. Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 40. <https://doi.org/10.1186/s41239-024-00468-z>
24. Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour*, 8(10), 1839–1850. <https://doi.org/10.1038/s41562-024-02004-5>
25. Zapata-Rivera, D., Torre, I., Lee, C.-S., Sarasa-Cabezuelo, A., Ghergulescu, I., & Libbrecht, P. (2024). Editorial: Generative AI in education. *Frontiers in Artificial Intelligence*, 7, 1532896. <https://doi.org/10.3389/frai.2024.1532896>
26. Zhou, X., Li, Y., Chai, C. S., & Chiu, T. K. F. (2025). Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 education from a systematic review. *Interactive Learning Environments*, 33(10), 5766–5788. <https://doi.org/10.1080/10494820.2025.2487538>



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non-Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.