

How Generative AI is used in Learning Games: A Systematic Scoping Review in K-12 Settings

Deniz Ercan, School of Teaching and Learning, CLASS Lab, University of Florida, deniz.ercan@ufl.edu
Seyedahmad Rahimi, CLASS Lab, College of Lifetime Learning, Georgia Tech., srahimi@gatech.edu

Abstract: This systematic scoping review examines the application of Generative Artificial Intelligence (GenAI) in K–12 STEM learning games. Following PRISMA guidelines, eight studies published between 2017 and 2025 were identified across six databases. Two main uses of GenAI emerged: learning support and assessment/analysis. Overall, Findings showed that GenAI supports learning through feedback and motivation but remains peripheral to core game mechanics.

Introduction

A persistent challenge in education is preparing students for a technology-driven world while sustaining engagement, particularly in K–12 STEM (NSB, 2023). NSB (2024) reports declining science scores, with only 31% of 8th graders reaching proficiency. Students often perceive STEM as abstract and disconnected, contributing to low engagement (Potvin & Hasni, 2014). Game-based learning (GBL) situates learning in interactive contexts where students actively explore and apply knowledge, supporting visualization of abstract concepts and learning by doing (Gee, 2003; Plass et al., 2015; Aldrich, 2005). Recent advances in generative AI (GenAI), including large language models, expand possibilities for personalized and adaptive learning (Bommasani et al., 2021; Zhai et al., 2021). GenAI can generate feedback, examples, and dialogue, potentially enhancing learning games through adaptive and interactive experiences (Chen et al., 2023; Cotton et al., 2024).

Despite rapid development, the use of GenAI in K–12 learning games remains underexplored. This review synthesizes existing research to clarify how GenAI has been applied and its role in learning processes.

Current study

This scoping review examines GenAI in K–12 STEM learning games. Prior reviews emphasize higher education or conceptual applications (Nzenwata et al., 2024; Yang et al., 2025), with limited focus on K–12 STEM contexts. The following research questions guide our scoping review:

- RQ 1) *What are the main trends in research on GenAI-powered learning games for K–12 students?*
- RQ 2) *What types of GenAI technologies have been used in GenAI-powered learning games for K–12 STEM, and in what ways have these technologies been applied?*

Method

This systematic scoping review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) to map how GenAI has been used in digital learning games in K–12 STEM contexts. Inclusion criteria were limited to full-text studies published in English between 2017 and 2025 that involved digital games, GenAI applications, STEM subjects, and K–12 mainstream students. Studies were excluded if they focused on non-digital games, non-GenAI applications, non-STEM subjects, higher education, adult learning, or review-type publications. The search was conducted across ERIC, ScienceDirect, ProQuest, Google Scholar, ACM Digital Library, and Scopus using combinations of terms related to game-based learning, generative AI, STEM education, and K–12 settings. One researcher reviewed titles and abstracts, followed by full-text screening of eligible studies using Covidence. After study selection, data were extracted and coded through an iterative process. One researcher developed the initial codes, and the second researcher reviewed and refined them. Through discussion, overlapping codes were merged into broader categories.

Results

Eight studies met the inclusion criteria. Two main categories emerged: learning support and assessment/analysis, with learning support focusing on feedback and motivation.

RQ 1: What are the main trends in research on GenAI-powered learning games for K–12 students?

The final studies comprised 8 studies, including journal articles ($n = 5$), conference papers ($n = 2$), and one book chapter ($n = 1$). Publication activity increased sharply in recent years: all studies ($n = 8$) were published in

2024–2025, reflecting growing interest following the public release of large GenAI models. Most research originated from the United States ($n = 5$), followed by Taiwan ($n = 2$) and China ($n = 1$). Science-focused games dominated ($n = 7$), with only one study addressing mathematics ($n = 1$). Role-playing games (RPGs) were the most common genre ($n = 5$), followed by adventure games ($n = 2$) and mini-games ($n = 2$). Most games incorporated interactive storytelling, dialogue-based problem solving, or inquiry-driven missions to scaffold conceptual understanding.

RQ 2) What types of AI technologies have been used in GenAI-powered learning games for K–12 STEM learning, and in what ways have these technologies been applied?

Based on the analysis of the 8 extracted studies, two major categories of GenAI applications in learning games were identified: (1) Learning Support and (2) Assessment/Analysis. Within these categories, learning support has two subcategories: 1) Feedback and 2) Motivation.

Learning support

In this category, most studies ($n = 6$) focused on feedback. Chen and Law (2025) and Huang and Chen (2025) integrated ChatGPT-based chatbots into science games to generate real-time explanations and problem-solving hints. Similarly, McLaren (2024) used GPT-3.5 within the *Decimal Point* game to interpret students' written self-explanations and generate rubric-based feedback, including identified misconceptions and suggested next steps. Goslen et al. (2025) fine-tuned T5 and GPT-3.5 models to produce goal-aligned problem-solving action plans from gameplay data. Wang et al. (2025) and Zhu et al. (2025) also examined conversational agents powered by Qwen and Llama 3 that interacted with players through dialogue and reflective questioning.

GenAI was also used as motivational learning support in two studies. Wang et al. (2025) and Zhu et al. (2025) designed agents that used supportive conversation, encouragement, and reflective prompts to sustain students' interest and perseverance. For example, the Qwen-powered agent “V” in *Planet Rosebud* provided personalized praise and questions when students hesitated or made errors. Similarly, Zhu et al. (2025) developed the peer agent *CAP*, which functioned as a companion rather than a tutor to maintain engagement during gameplay.

Assessment/Analysis

A smaller but growing body of work ($n = 3$) explored GenAI for assessment or analysis in learning games. These studies used GenAI to interpret or evaluate students' in-game reasoning and actions. Acosta et al. (2025) fine-tuned a T5 model to recognize epistemic dialogue acts in students' collaborative discussions, identifying moments of reasoning and knowledge exchange during gameplay. Liu et al. (2025) combined GPT-4o with an XGBoost classifier to detect struggle behaviors in gameplay by converting gameplay logs into text replays readable by the language model. McLaren (2024) also used GPT-3.5 to evaluate the quality of students' reasoning in self-explanations as correct, partially correct, or incorrect. Overall, these studies show that GenAI can be used to assess students' reasoning and gameplay behaviors and provide insights into learning processes and in-game performance.

Discussion and conclusion

This systematic scoping review mapped the emerging use of GenAI in digital learning games for K–12 STEM education. The studies reviewed suggest a gradual shift in game design from rule-based or scripted systems to adaptive, conversational, and data-driven applications of GenAI. Studies used conversational agents' responses in real time (Chen & Law, 2025; McLaren, 2024), yet their integration often remains peripheral to core gameplay. This suggests that current implementations use GenAI to supplement instruction rather than to drive gameplay mechanics. Embedding GenAI into game logic would enable dynamic content, adaptive tasks, and narrative adjustments based on learner progress. As part of the game engine, GenAI could create adaptive scenarios that respond to learners' behavior and sustain engagement. Such integration could transform GenAI from a feedback provider into an active co-designer of the learning experience. From an instructional perspective, the reviewed studies showed how GenAI can mediate both feedback and motivational support. These applications demonstrate GenAI's capacity to act as a learning partner that motivates students and fosters their engagement with game tasks. Yet, the interaction patterns with GenAI across studies remain primarily reactive, not proactive. Overall, GenAI supports interaction and analysis in learning games but remains underused as a generative mechanism. Future research should explore integrating GenAI into core design processes to create more adaptive and responsive learning experiences.

References

- Acosta, H., Lee, S., Bae, H., Feng, C., Rowe, J., Glazewski, K., ... & C. Lester, J. (2025). Recognizing multi-party epistemic dialogue acts during collaborative game-based learning using large language models. *International Journal of Artificial Intelligence in Education*, 35(2), 677-701.
- Al Braiki, B., Harous, S., Zaki, N., & Alnajjar, F. (2020). Artificial intelligence in education and assessment methods. *Bulletin of Electrical Engineering and Informatics*, 9(5), 1998-2007.
- Aldrich, C. (2005). Learning by doing: A comprehensive guide to simulations, computer games, and pedagogy in e-learning and other educational experiences. John Wiley & Sons.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Ahn, L., ... Liang, P. (2021). On the opportunities and risks of foundation models. Stanford Center for Research on Foundation Models.
- Chen, C., Zhou, H., & Yu, T. (2023). Game-based features in intelligent game-based learning environments: A systematic review. *Interactive Learning Environments*.
- Chen, C. H., & Law, V. (2025). The role of help-seeking from ChatGPT in digital game-based learning. *Educational technology research and development*, 1-19. <https://doi.org/10.1007/s11423-025-10464-3>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.
- Goslen, A., Kim, Y. J., Rowe, J., & Lester, J. (2025). Llm-based student plan generation for adaptive scaffolding in game-based learning environments. *International journal of artificial intelligence in education*, 35(2), 533-558. <https://doi.org/10.1007/s40593-024-00421-1>
- Huang, K., & Chen, C. H. (2025). Instructional Video and GenAI-Supported Chatbot in Digital Game-Based Learning: Influences on Science Learning, Cognitive Load and Game Behaviours. *Journal of Computer Assisted Learning*, 41(4), e70094. <https://doi.org/10.1111/jcal.70094>
- Liu, X., Wei, Z., Baker, R. S., Metcalf, S. J., Zhang, J., Barany, A., ... & Gagnon, D. J. (2025, July). Integrating large language models and machine learning to detect struggle in educational games. In *International Conference on Artificial Intelligence in Education* (pp. 398-405). Cham: Springer Nature Switzerland.
- McLaren, B. M. (2024). Decimal point: a decade of learning science findings with a digital learning game. In *Artificial Intelligence in Education: The Intersection of Technology and Pedagogy* (pp. 145-203). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-71232-6_9
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Bmj*, 339. <https://doi.org/10.1136/bmj.b2535>
- National Science Board. (2024). *The STEM labor force: Scientists, engineers, and skilled technical workers* (NSB-2024-5). <https://nces.nsf.gov/pubs/nsb20245>
- National Science Board. (2023). *Elementary and secondary STEM education*. In *Science and Engineering Indicators 2024* (NSB-2023-31). National Science Foundation. <https://nces.nsf.gov/pubs/nsb202331>
- Nzenwata, U. J., CL, B. N., Ojelabi, E. O., Oduware, O., Yisau, Y., Adeyela, A. T., ... & Osisanya, O. A. (2024). A systematic review of generative AI in education. *Journal of Computer Sciences and Applications*, 12(1), 25-30. <https://doi.org/10.12691/jcsa-12-1-4>
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational psychologist*, 50(4), 258-283.
- Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85-129.
- Tolks, D., Schmidt, J. J., & Kuhn, S. (2024). The role of AI in serious games and gamification for health: scoping review. *JMIR Serious Games*, 12(1), e48258. <https://doi.org/10.2196/48258>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Wang, M., Zhang, D., Zhu, J., & Gu, H. (2025). Effects of Incorporating a Large Language Model-Based Adaptive Mechanism Into Contextual Games on Students' Academic Performance, Flow Experience, Cognitive Load and Behavioral Patterns. *Journal of Educational Computing Research*, 63(3), 662-694.
- Yang, D., Kleinman, E., & Hartevel, C. (2025). GPT for Games: An Updated Scoping Review (2020-2024). *IEEE Transactions on Games*. <https://doi.org/10.1109/TG.2025.3563780>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021(1),
- Zhu, C., Sam, C. H., Wu, Y., & Tang, Y. (2025). Wip: Enhancing game-based learning with ai-driven peer agents. *arXiv preprint arXiv:2508.01169*. <https://doi.org/10.48550/arXiv.2508.01169>