

Generative AI for Creative Learning in K-12 Education: Insights from a Systematic Scoping Review

Seyedahmad Rahimi, CLASS Lab, College of Lifetime Learning, Georgia Tech., srahimi@gatech.edu

Maryam Babae, School of Teaching and Learning, CLASS Lab, University of Florida,

babae.maryam@ufl.edu

Salah Esmailigoujar, School of Teaching and Learning, CLASS Lab, University of Florida,

s.esmailigoujar@ufl.edu

Chris Dede, Harvard Graduate School of Education, chris_dede@gse.harvard.edu

Abstract: The rapid rise of generative AI (GenAI) presents both opportunities and challenges for fostering creativity in K–12 education. This systematic scoping review synthesized 45 studies (2017–2025) to examine how GenAI has been applied in creative learning. Four categories emerged: (1) creativity assessment, (2) human–AI co-creativity, (3) stakeholders’ perceptions, and (4) creativity enhancement. Findings show strong potential for GenAI to support idea generation, multimodal expression, and assessment of creative outputs, but highlight concerns about equity, authorship, and theoretical grounding. While most applications focus on text-based creativity, multimodal and culturally inclusive approaches remain underexplored. This review provides a foundation for responsible and equitable integration of GenAI into K–12 classrooms and captures a field in transition, where new forms of creative expression are emerging and educators are rethinking assessment, authorship, and pedagogy. This review provides a foundation for ethically grounded, developmentally appropriate and theory-driven GenAI designs that enhance, rather than automate, human creativity.

Introduction

Creativity is central to preparing young learners for the demands of the 21st century, yet it has been historically marginalized in K–12 education, where standardized testing and rote memorization dominate (Craft, 2006; OECD, 2022). Creativity is defined as the ability to go beyond established norms to generate original and meaningful ideas which forms the foundation of innovation and social progress (Amabile, 1983; Kaufman & Beghetto, 2009). Supporting creativity from early childhood builds learners’ curiosity, resilience, and critical thinking (Kupers et al., 2019). Despite its recognized value, educators often have lacked resources or training to foster or measure creativity, and have often relied on subjective, labor-intensive assessments (Blamires & Peterson, 2014).

GenAI tools such as large language models (LLMs) and text-to-image generators are now entering classrooms. These tools offer the potential to scaffold ideation, personalize feedback, and automate aspects of creativity assessment (Wingstrom et al., 2024). However, irresponsible use risks plagiarism, homogenization, and bias (Hagerty & Rubinov, 2019; Shah & Sureja, 2025). Unlike earlier AI systems that focused on prediction, personalization, and analytics (Vázquez-Cano et al., 2021), GenAI shifts the focus toward co-creation or automation, which enables learners to engage in human-AI partnerships that combine algorithmic power with human imagination, where technology becomes a thinking partner rather than a passive instrument (Wingstrom et al., 2024). GenAI can also support creativity through real-time multimodal feedback and more authentic assessments such as stealth assessment, which unobtrusively evaluates creativity during digital or game-based tasks (Shute, 2011; Rahimi & Shute, 2023). This shift demands not only technical fluency but also a renewed pedagogical understanding of creativity as a social, ethical, and contextual process.

At the same time, this partnership demands ethical literacy; teachers and students must be able to identify algorithmic bias, representation gaps, and the limits of AI-generated work (Hagerty & Rubinov, 2019). These challenges align with the five-cycle ethical framework (Dieterle et al., 2024), which identifies divides in access, representation, algorithms, interpretation, and citizenship as key areas influencing responsible AI use in schools. Developing this awareness is essential to ensure GenAI serves as a tool for amplifying human creativity rather than diminishing it. By linking GenAI’s creative potential with such an ethical framework, educators can design learning environments that are equitable, transparent, and human-centered (see Rahimi et al., 2026).

This paper reports on a scoping review of studies published between 2017–2025 that examined the role of GenAI in supporting or assessing creativity in K–12 contexts. Our goal was to map trends, identify opportunities, and surface limitations to inform equitable, theory-informed, and ethically grounded integration of GenAI into schools (Dieterle et al., 2024; Rahimi, 2023). Ultimately, this review aims to guide future practices that use GenAI not to automate creativity, but to augment it in ways that empower both teachers and learners.

Method

A systematic scoping review approach was chosen to capture the breadth of emerging evidence in this rapidly evolving field, which allows for diverse study designs and exploratory synthesis (Peters et al., 2015). Following PRISMA guidelines (Liberati et al., 2009), we searched six academic databases (Web of Science, ERIC, Scopus, ScienceDirect, ProQuest, and Google Scholar) and key creativity and AI in Education journals for studies published between 2017–2025. We limited the time window to 2017–2025, because 2017 marks the introduction of transformer architectures that underpin modern generative models (Vaswani et al., 2017).

Our main search focused on combinations of terms “creativity,” “generative AI,” “K-12,” and included related terms such as LLMs, and ChatGPT. We also conducted hand searches of reference lists and the publications of leading scholars. Because databases differ in syntax and filters, we adapted the search strategy for each one. Using broad and inclusive keywords increased the likelihood of capturing all relevant studies across platforms. Inclusion and exclusion criteria are presented below (Table 1). In total, 45 studies met the criteria and were thematically analyzed by a team of four researchers (Figure 1).

Table 1

Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Published 2017–2025	Published before 2017
English language	Non-English
Focus on K–12 education	Focus on adult/higher ed or informal settings
Creativity as a primary or secondary outcome	No mention of creativity
Use of GenAI (e.g., LLMs, text-to-image tools)	Traditional or rule-based AI only
Full-text access	No full-text available

Results

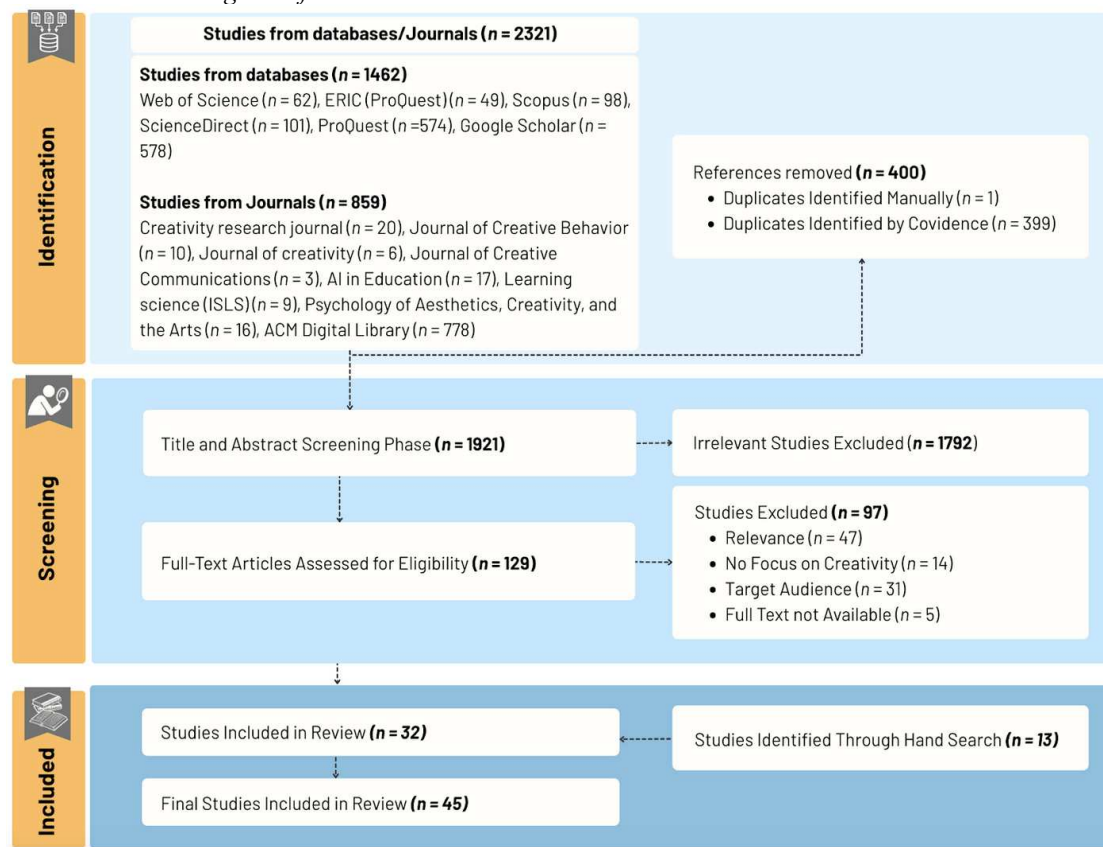
We retrieved 2,321 records and removed 399 duplicates. During title and abstract screening, one researcher conducted the review, and 129 full-text articles advanced to full screening by two researchers using the inclusion criteria. To strengthen coverage, we also conducted a targeted hand-search of the recent work of key researchers whose work lies at the intersection of creativity, AI, and education, using Google Scholar, ResearchGate, and conference proceedings. This process yielded 13 additional studies. In total, 45 articles were included and synthesized thematically (see Figure 1).

Inductive coding was used by three researchers working independently to identify emerging themes within the selected studies. After the initial coding, researchers met and resolved conflicts. Thematic patterns were validated through team triangulation, which ensures that categories accurately reflected the diversity of the reviewed studies.

Across the reviewed studies, quantitative and mixed-methods designs were the most common. Most publications appeared between 2022–2025, which shows a rise in interest following the public release of large generative models such as ChatGPT.

In our review, we identified four overarching categories of research: (1) *Creativity Assessment*; (2) *Stakeholders’ Perceptions*; (3) *Human–AI Co-Creativity*; (4) *Creativity Enhancement*. In the first category, *Creativity Assessment* ($n = 4$), studies examined the use of GenAI-based tools to assess divergent thinking and creative products. These studies included MOTES, a computational psychometric tool for originality (Acar et al., 2023), a multimodal model for scientific creativity scoring (Goecke et al., 2024), GPT-4o applied to assessing game-level creativity in *Physics Playground* (Rahimi, 2024), and a conceptual synthesis of an AI-based creativity assessment framework (Acar, 2025). Results showed strong correlation between AI-generated and expert ratings, suggesting potential for scalable assessment, while raising concerns about bias, validity, and transparency persist. Across the four studies, GenAI-based scoring aligned well with human expert ratings. It offers a potential path for scaling creativity assessment, though its accuracy and fairness still need careful attention. Moreover, these studies point to a shift toward hybrid assessment models where AI works with, not instead of human judgment. By capturing patterns of originality and elaboration across diverse modalities such as text, visuals, and game play, GenAI expands what can be measured as creativity. Although, studies point out that we should be cautious about how these measures are interpreted and applied.

Figure 1
PRISMA Flow Diagram of Studies Selection



In the second category, *Stakeholders' Perceptions* ($n = 14$), studies captured how students, teachers, parents, and designers perceive GenAI role in creativity. Stakeholders generally viewed GenAI as a scaffold that supports rather than replaces human creativity. Across studies, participants emphasized four recurring priorities: (1) cultural representation; (2) emotional resonance; (3) transparency; and (4) maintaining student agency as critical for meaningful integration. Stakeholders often described creativity as a social and cultural practice shaped by lived experiences and emotional investment (Han & Cai, 2023). Teachers and designers stressed the need for culturally responsive AI design, while students emphasized trust and authorship as prerequisites for engagement. These findings underscore that emotional connection and transparency are vital to sustaining creative ownership when learners use GenAI tools. Moreover, stakeholders frequently described their engagement with GenAI as a process of negotiations rather than acceptance. They valued the inspiration and efficiency that AI could offer but remained alert to the risk of overreliance and loss of personal expression.

In the third category, *Human-AI Co-Creativity* ($n = 3$), studies explored how learners collaborate with GenAI agents during open-ended creative tasks such as storytelling and writing. When GenAI acted as a responsive partner rather than a directive tool, students exhibited higher idea fluency, engagement, and creative confidence. Co-creative designs that allowed turn-taking, elaboration, and narrative feedback promoted sustained creativity and enjoyment (Ding et al., 2024). Two patterns emerged from these studies. First, students with higher baseline creativity and more positive views of AI reported greater trust and were more likely to continue using the system (Ding et al., 2024). Second, designs that made AI contributions easy to inspect and revise—for example editable suggestions and brief reflection prompts—helped preserve authorship and reduce overreliance. However, overreliance on AI, limited personalization, and the lack of social presence or emotional depth were noted as barriers to effective collaboration. These findings suggest that productive co-creativity depends on balancing system autonomy with human agency and transparency.

In the fourth and largest category, *Creativity Enhancement* ($n = 34$), studies examined how GenAI tools are used to foster creative thinking and learning. This category included three subgroups: (1) Creativity Support Tools ($n = 11$), where GenAI tools were intentionally designed or adapted to scaffold creative processes; (2) Specific Effects of GenAI on Creativity ($n = 8$), where creativity was the central focus of instruction or

intervention; and (3) Broad Effects of GenAI on Creativity ($n = 15$), where creativity emerged as an outcome of GenAI-supported learning experiences.

Across these studies, storytelling and creative writing tasks were most prevalent, highlighting the dominance of language-based creativity in K-12 research. Fewer studies addressed auditory, embodied, or spatial creativity, pointing to the limited exploration of multimodal expression. Moreover, few explicitly drew on established creativity theories such as the 4C Model (Kaufman & Beghetto, 2009) or Csikszentmihalyi's Systems Theory (Csikszentmihalyi, 1999), which limits conceptual depth and comparability across studies. Overall, this imbalance suggests that while GenAI is being integrated widely, it is not yet being studied with consistent theoretical depth or methodological rigor.

Discussion and conclusion

This systematic scoping review reveals that while the rapid rise of GenAI offers promising avenues to scale and support creativity in K-12 education, its value depends on intentional, ethical, and theory-driven design. Across four main categories, Creativity Assessment, Stakeholders' Perceptions with four sub-categories: Creativity as a Social and Cultural Practice; Emotional Resonance and Personal Meaning; Transparency, Explainability, and Trust; and The Centrality of Student Agency and Creative Ownership, Human-AI Co-Creativity, and Creativity Enhancement with three sub-categories: Creativity Support Tools, Specific Effects of GenAI on Creativity, and Broad Effects of GenAI on Creativity.

The reviewed studies (2017 to 2025), indicates a rapid rise in search on GenAI's role in creative learning. The findings highlight GenAI's potential and its constraints as both a creative support and assessment tool in K-12 settings. Researchers are beginning to position GenAI as a catalyst for idea generation, multimodal expression, and inclusive creative learning. These findings align with the border vision of creativity augmentation and intelligence augmentation (Dede et al., 2021; Rahimi et al., 2026), which emphasizes GenAI as a collaborator that amplifies rather than replaces learners' agency. In practice, this suggests that AI systems should be designed to prompt reflection, dialogue, and critique, not just generate content, which deepens the educational value of creative engagement.

A central finding across studies is that GenAI's influence is contextual rather than automatic. Tools add value when they (a) make reasoning visible and contestable for learners and teachers, (b) preserve authorship and voice, (c) are tuned to developmental, cultural, and disciplinary contexts. In other words, the pedagogical wrapper, transparent rubrics, reflective prompts, culturally responsive examples, and teacher mediation matter as much as the model itself. Where these conditions were present, studies reported stronger engagement, richer elaboration, and more sustained creative exploration.

At the same time, the field would benefit from firmer theoretical anchoring. Many studies use creativity as mere novelty or idea fluency but under-specify process, motivation, and sociocultural contexts. Greater use of established frameworks such as the 4C Model (Kaufman & Beghetto, 2009) or Csikszentmihalyi's Systems Theory (Csikszentmihalyi, 1999), would improve construct clarity and comparability across studies. Theory-informed designs also help researchers specify *which* facet of creativity is meant to cultivate and *how* learning environments should be structured to get there. This theoretical grounding can also serve as a bridge between creativity assessment and creativity enhancement, ensuring coherence across how creativity is measured, taught, and supported.

Ethically, GenAI in creative learning must be situated within an explicit equity lens. Dieterle et al.'s (2024) five-cycle framework (i.e., access, representation, algorithms, interpretation, and citizenship) offers a practical scaffold. Doing so would shift evaluation from "Does it work?" to "For whom, under what conditions, and with what consequences?" This reframe is essential for protecting authorship, avoiding representational harm, and ensuring that students learn to critique, revise, and *own* AI-mediated outputs. Such reflection is especially urgent as schools increasingly rely on commercial AI systems that may reproduce systematic inequities or marginalized non-dominant voices.

This study has important implications for educators and policymakers seeking to integrate GenAI responsibly in K-12 classrooms. Professional development should prepare teachers to address issues of bias, transparency, and authorship while using GenAI to promote student agency and creativity. Students, in turn, need AI literacy programs that build both critical understanding and creative confidence. Clear classroom policies and ongoing evaluation can ensure that GenAI strengthens, rather than replaces, the human dimensions of creative learning.

Future research should move toward designing and evaluating GenAI systems that are transparent, interpretable, and guided by creativity theory. Collaboration between educators, AI developers, and policymakers will be essential to ensure that these tools promote inclusivity and safeguard learner agency.

References

- Acar, S., Dumas, D., Organisciak, P., & Berthiaume, K. (2024). Measuring original thinking in elementary school: Development and validation of a computational psychometric approach. *Journal of Educational Psychology*. <https://psycnet.apa.org/doi/10.1037/edu0000844>
- Acar, S. (2025). Creativity assessment, research, and practice in the age of artificial intelligence. *Creativity Research Journal*, 37(2), 181–187.
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, 45(2), 357.
- Blamires, M., & Peterson, A. (2014). Can creativity be assessed? Towards an evidence-informed framework for assessing and planning progress in creativity. *Cambridge Journal of Education*, 44(2), 147–162. <https://doi.org/10.1080/0305764X.2013.860081>
- Craft, A. (2006). Fostering creativity with wisdom. *Cambridge Journal of Education*, 36(3), 337–350.
- Csikszentmihalyi, M. (1999). Implications of a systems perspective for the study of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 313–335). Cambridge University Press.
- Dede, C., Etemadi, A., & Forshaw, T. (2021). *Intelligence augmentation: Upskilling humans to complement AI*. The Next Level Lab at the Harvard Graduate School of Education. President and Fellows of Harvard College: Cambridge, MA.
- Ding, Z. Y., & Chiu, M. H. (2024). Can I collaborate with you? An investigation of trust in generative AI. *Proceedings of the Association for Information Science and Technology*, 61(1), 889–891.
- Dieterle, E., Dede, C., & Walker, M. (2024). The cyclical ethical effects of using artificial intelligence in education. *AI & Society*, 39(2), 633–643.
- Goecke, B., DiStefano, P. V., Aschauer, W., Haim, K., Beaty, R., & Forthmann, B. (2024). Automated scoring of scientific creativity in German. *The Journal of Creative Behavior*, 58(3), 321–327. <https://doi.org/10.1002/jocb.658>
- Han, A., & Cai, Z. (2023, June). Design implications of generative AI systems for visual storytelling for young learners. In *Proceedings of the 22nd annual ACM interaction design and children conference* (pp. 470–474). <https://doi.org/10.1145/3585088.3593867>
- Hagerty, A., & Rubinov, I. (2019). Global AI ethics: a review of the social impacts and ethical implications of artificial intelligence. *arXiv preprint arXiv:1907.07892*.
- Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of general psychology*, 13(1), 1–12. <https://doi.org/10.1037/a0013688>
- Kupers, E., Lehmann-Wermser, A., McPherson, G., & van Geert, P. (2019). Children's creativity: A theoretical framework and systematic review. *Review of Educational Research*, 89(1), 93–124.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of internal medicine*, 151(4), W-65.
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Baldini Soares, C. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141–146. <https://doi.org/10.1097/XEB.0000000000000050>
- Rahimi, S., & Shute, V. J. (2023). Stealth assessment: A theoretically grounded and psychometrically sound method to assess, support, and investigate learning in technology-rich environments. *Educational Technology Research and Development*, 72, 2417–2441. <https://doi.org/10.1007/s11423-023-10232-1>
- Rahimi, S., Dede, C., Esmailigoujar, S., & Babaee, M. (2026). Augmenting human creativity with responsible and ethical use of generative AI. In M. J. Worwood & J. C. Kaufman (Eds.), *Generative artificial intelligence and creativity (Explorations in Creativity Research)*, pp. 87–99. Academic Press.
- Shah, M., & Sureja, N. (2025). A comprehensive review of bias in deep learning models: Methods, impacts, and future directions. *Archives of Computational Methods in Engineering*, 32(1), 255–267.
- Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). IAP Information Age Publishing.
- Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18(33). <https://doi.org/10.1186/s41239-021-00269-8>
- 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
- Wingström, R., Hautala, J., & Lundman, R. (2024). Redefining creativity in the era of AI? Perspectives of computer scientists and new media artists. *Creativity Research Journal*. Advance online publication. <https://doi.org/10.1080/10400419.2022>