

Repositioning STEM Education in the Context of Generative Artificial Intelligence: A Review Focusing on Critical Thinking and Problem Solving

Lin Peng*

Guangzhou Maritime University, Guangdong, China

**Corresponding author: Lin Peng.*

Abstract

In recent years, a wealth of studies have explored the integration of generative artificial intelligence (AI) into classroom teaching within China's education sector. To clarify the overall biases embedded in existing scholarship, this review sorts out Chinese and English journal articles published between 2022 and 2026, and compares divergent research priorities between Chinese and international academic communities. Most domestic studies on generative AI-enabled STEM education center on developing intelligent teaching tools and constructing digital classroom environments, while only a handful of top-tier theoretical discussions address the reconstruction of educational objectives and the cultivation of higher-order thinking competencies. Collectively, current research exhibits a pronounced bias that prioritizes technological advancement over the core essence of education. Against this notable research gap, the present study identifies critical thinking and problem-solving capabilities as the two core educational objectives for STEM classrooms. It further proposes actionable implementation strategies across four dimensions: classroom instruction, project design, assessment systems, and AI usage governance, offering practical references for digital STEM teaching in primary, secondary and tertiary institutions nationwide. This review is structured around five core components: prevailing pain points in real-world classroom practice, gaps in the extant literature, literature sorting methods adopted, key research findings, and theoretical as well as instructional implications.

Keywords

generative artificial intelligence, STEM education, educational repositioning, critical thinking, problem-solving competency

1. Introduction

Competency-based education has long guided global education reform. For years, the Programme for International Student Assessment (PISA) framework has underscored the educational value of creative thinking and integrated problem-solving skills. The Organisation for Economic Co-operation and Development (OECD) formally defines creative thinking in its 2021 PISA assessment report: Creative thinking refers to the capacity to productively generate, evaluate and refine ideas that yield original, effective solutions [1]. Critical thinking and problem-solving are likewise stipulated as core training outcomes for STEM disciplines in China's national compulsory education curriculum standards. Nevertheless, on-site classroom observations reveal that, driven by examination pressure, most schools remain focused on rote knowledge

memorization and standardized drill exercises, leaving little room for authentic inquiry-based learning. Field research conducted by Yang confirms that 82.7% of STEM classrooms in China's primary and secondary schools still prioritize knowledge delivery and standardized training, with inquiry-oriented thinking activities accounting for an extremely small proportion of instruction, markedly departing from the interdisciplinary mission of STEM education [2]. The widespread accessibility of GPT-series large language models has drastically lowered the technical barrier for students to generate problem-solving procedures and complete written assignments. The official GPT-4 technical report published by OpenAI states: GPT-4 is a multimodal large language model capable of processing both image and text inputs, autonomously generating mathematical derivations, engineering simulations and full academic manuscripts with minimal operational complexity [3]. Cross-regional questionnaire surveys document rampant misuse of such tools: 63% of middle school students and 71% of university undergraduates rely entirely on generative AI to finish after-class homework. Nearly half of learners abandon independent logical deduction and directly copy AI-generated answers. Statistical comparisons show that students who habitually leverage intelligent tools to complete logical assessments score 18.6% lower on average than peers who finish tasks independently [1, 4]. Frontline teaching experience also indicates a growing tendency among students to seek instant final results from AI, accompanied by a steady decline in willingness to engage in independent exploratory reasoning. Academic misconduct involving AI-assisted assembly of experimental reports and course design documents has risen year on year. Faced with these persistent practical challenges, it becomes imperative to redefine the educational positioning of STEM and restructure classroom delivery and evaluation mechanisms accordingly.

This review screening exclusively includes peer-reviewed journal publications in Chinese and English; conference proceedings and unpublished dissertations are excluded from the sample. As such, the resulting analytical findings only reflect general industry trends and carry inherent limitations. During the literature collation process, this review addresses three central research questions: First, what pedagogical benefits and cognitive risks does generative AI bring to STEM classrooms? Second, what divides exist between domestic and international scholarly research agendas, and what critical gaps remain unaddressed in current literature? Third, how can a complete AI-adaptive STEM teaching framework be constructed centering on the two aforementioned higher-order competencies? Section 2 defines core conceptual terms and elaborates the foundational theoretical frameworks supporting this research. Section 3 draws on multiple empirical datasets to discuss both the opportunities and persistent pedagogical hurdles introduced by generative AI (this section streamlines discussion of secondary phenomena to shorten its length). Section 4 employs bibliometric data to compare domestic and international research trajectories, with a focused analysis of deficiencies within Chinese scholarship. Section 5 justifies the necessity of STEM educational repositioning from three standpoints: theoretical foundations, contemporary technological context, and real classroom intervention experiments. Section 6 details feasible implementation strategies for the four dimensions of AI-enabled STEM instruction. The final section summarizes the paper's core arguments, acknowledges inherent limitations as an undergraduate literature review, and outlines preliminary directions for future related inquiry.

2. Theoretical Basis

2.1 Definition of Core Concepts

2.1.1 Generative Artificial Intelligence

Conventional intelligent software is limited to information retrieval and basic numerical computation, fundamentally differing from generative AI systems. OpenAI's technical documentation distinguishes traditional and emerging intelligent paradigms: Unlike analytical AI constrained to retrieval and arithmetic tasks, generative AI learns from massive training datasets to produce original textual content, mathematical models and engineering blueprints [3]. Such user-friendly tools support full-scenario application including pre-class preview, in-class exploration and after-school homework, yet they carry both merits and drawbacks. He put forward the classic definition of scaffolding instruction: "Scaffolding teaching breaks down complex tasks through layered prompts, builds cognitive steps, guides students to think independently, and never directly provides complete answers" [9]. If students form the habit of copying AI-generated content, the whole process of self-analysis, trial and error, and reflective thinking will be completely eliminated. Zimmerman illustrated the core logic of self-regulated learning: Self-regulated learning relies on continuous self-examination and logical reflection, which will be severely inhibited once students directly copy outputs from AI tools [4].

Neuroscience experimental evidence further proves that long-term dependence on AI for ready-made answers leads to markedly weakened activity in brain regions responsible for logical reasoning. This demonstrates that clear boundaries for AI tool usage must be defined in classroom teaching based on human-machine collaborative learning theory.

2.1.2 STEM Education

Many teachers hold a biased superficial understanding that STEM merely pieces together knowledge points across four disciplines. According to empirical research by Yang, STEM is not a simple combination of subject knowledge, but takes authentic, complicated engineering problems as the carrier to lead students to conduct exploration, design and iterative optimization with multidisciplinary knowledge [2]. Large-scale nationwide classroom surveys reveal that 82.7% of primary and secondary STEM classrooms prioritize theoretical lectures and standardized experimental operations, lacking inquiry-oriented training and deviating from the original educational mission of STEM [2]. From the perspective of social constructivism, learning activities depend on student collaboration and real-life contexts to build personal independent cognition, naturally cultivating high-order abilities such as innovation and comprehensive problem-solving, abilities that precisely match the talent demands of the AI era.

2.1.3 Critical Thinking and Problem-Solving Abilities

Critical thinking serves as the core competency enabling students to identify errors in AI outputs. Zimmerman defined the complete process of critical thinking as follows: Critical thinking refers to examining evidence, conducting logical analysis and evaluating external information from multiple perspectives [4]. Full critical thinking training involves evidence verification, logical sorting and multi-dimensional assessment, which effectively avoids learning biases and misunderstandings caused by one-sided or false AI-generated content. Problem-solving runs through all STEM inquiry activities, covering the full cycle of problem identification, goal decomposition, solution design and iterative optimization. The official OECD report notes: Basic memory and calculation work can be automated, while complex problem-solving and critical judgment remain exclusive human strengths [1]. Multiple classroom control experiments also confirm that algorithms cannot yet replicate human comprehensive thinking and creative capacity. For this reason, this review takes these two capabilities as core training objectives of STEM education.

2.2 Relevant Theoretical Framework

2.2.1 Cultivation Theory of STEM Higher-Order Thinking

Zimmerman's classic exposition on self-regulated learning states: Self-regulated learning is a cyclic process consisting of hypothesis formulation, inquiry, verification and reflection [4]. This viewpoint aligns closely with the core tenet of social constructivism: knowledge cannot be unilaterally instilled by teachers, but must be actively constructed by learners themselves. Inquiry-based learning further refines classroom exploration into a closed loop from hypothesis raising to empirical verification, continuously exercising students' logical reasoning and innovative thinking. The integration of the two theories underpins this review's reform proposal of "reducing mechanical computational drills and focusing on high-order competency cultivation", and acts as the theoretical foundation for subsequent classroom design schemes.

2.2.2 Theory of Human-Computer Collaborative Learning

Human-machine collaborative learning theory clearly divides instructional responsibilities among teachers, students and generative AI. Domestic scholars Fang and Kong summarized the division-of-labor principle in their research: "Artificial intelligence undertakes low-level repetitive tasks including information sorting and basic calculation, while teachers and students concentrate on high-order cognitive activities such as logical analysis and innovative design" [8]. AI functions only as an auxiliary cognitive tool and can never replace the dominant learning position of students. Classroom adjustment guided by this labor division can effectively ease the problem of excessive student reliance on intelligent tools. The initiative of classroom exploration and independent thinking always belongs to students, and this theory runs through all subsequent designs of teaching activities, assessment mechanisms and risk prevention measures.

3. The Dual Impact of Generative AI on STEM Education

3.1 Opportunities for Educational Development

Generative AI can optimize existing STEM classroom models in three dimensions: personalized tutoring, interdisciplinary project delivery and differentiated layered teaching. Chen and Wang verified through controlled classroom experiments that “STEM classrooms adopting properly designed AI-assisted scaffolding tools can raise student participation in inquiry activities by 26.3%, and greatly lower the implementation threshold of interdisciplinary projects” [6]. When carrying out engineering exploration projects, repetitive work such as data statistics and model simulation can be assigned to AI, freeing teachers and students to focus on discussing innovative solutions. Meanwhile, intelligent tools can push customized learning materials matching students’ proficiency levels, satisfying the consolidation demands of underachieving learners and the expansion needs of advanced students simultaneously.

3.2 Realistic Challenges Faced by Traditional Teaching

Conclusions drawn from electroencephalogram (EEG) experiments show that frequent direct copying of AI outputs suppresses the activity of logical reasoning brain regions by 31% [1]. Long-term reliance on ready-made AI answers fosters rigid thinking inertia; students only care about final results while ignoring deductive reasoning procedures, resulting in superficial pseudo-learning. Literature sorting conducted by Wu et al. finds that academic misconduct involving AI-assisted completion of experimental reports and course design documents has risen by 47% compared with 2021, following the popularization of generative large models [7]. Traditional evaluation systems relying on written examinations and homework grading can no longer objectively measure students’ real thinking levels, and the conventional teaching mode centered on knowledge memorization and computational training is unsustainable under the new technological context.

3.3 The Necessity of Repositioning STEM Education

AI has already taken over low-level learning tasks such as knowledge memorization and basic mathematical computation. If STEM classrooms still prioritize mechanical problem drills and knowledge recitation, they will lose their unique educational value. The OECD report confirms that critical thinking and complex problem-solving constitute humanity’s irreplaceable competitive advantages [1]. Combined with social constructivism, human-machine collaborative learning theory and abundant domestic and international classroom intervention experiment results [6], it is clear that repositioning STEM education is an inevitable choice to adapt to technological evolution and return to the intrinsic essence of education.

4. Analysis of Domestic and Foreign Research Status

4.1 Development Trends of Foreign Research

Overseas research on generative AI in STEM education started earlier with a clear three-stage evolutionary trajectory [6]. The 2018–2021 period marks the tool testing stage, where scholars mainly verified the classroom application efficiency of various AI software tools. After the launch of large multimodal models in 2022, academia began reflecting on weakened student thinking ability and academic fraud triggered by tool abuse. Since 2024, research priorities have shifted to practical themes including human-machine collaboration and classroom cultivation of critical thinking. Bibliometric data sorted by Zheng and Zhang indicates that 68% of overseas English-language papers set high-order competencies as core research variables, supported by abundant frontline controlled classroom experiments [10]. International research has completed the paradigm shift from tool efficiency testing to ability cultivation research, with more mature integrated theoretical and practical systems.

4.2 Domestic Research Status

Domestic research closely follows the iteration pace of intelligent technology, yet notable gaps exist between its research focus and international scholarship. According to original statistical data compiled by Zheng and Zhang, “72.4% of domestic relevant publications focus on technological development such as intelligent teaching assistants and online assessment systems, while merely 11.3% of papers discuss top-level

design of educational objectives and high-order thinking cultivation” [10]. A review of recent domestic journal articles demonstrates a severe scarcity of localized complete theoretical frameworks centered on the two core high-order competencies proposed in this review. Controlled empirical classroom intervention studies remain relatively scarce, and domestic research generally prioritizes tool development while neglecting the construction of educational core values.

4.3 Research Differences and Innovation Points

The fundamental root of divergent domestic and international research orientations lies in distinct starting points of academic exploration: overseas researchers take STEM education itself as the core research object, prioritizing overall paradigm reform of classroom instruction and educational objectives; most domestic studies focus on software functional development and specific classroom operation methods, lacking in-depth macro-level educational reflection, which explains the insufficient depth of relevant domestic research. Targeting multiple deficiencies identified in extant literature, this review puts forward three innovative perspectives: First, breaking the single-dimensional research perspective limited to tool application adopted by most existing papers, and conducting analysis from the macro perspective of reconstructing top-level educational objectives. Second, integrating two mature theories (social constructivism and human-machine collaborative learning) to build a localized dual-core theoretical framework for domestic STEM education practice. Third, designing operable classroom implementation strategies matching the actual teaching conditions of China’s primary, secondary and tertiary schools. As a literature review finished by an undergraduate student, this review only deduces conclusions based on published public literature without access to on-site comparative controlled classroom experiments in schools. Accordingly, the innovative viewpoints proposed herein still require verification through follow-up empirical research, which constitutes a major limitation of this study.

5. The Core Repositioning of STEM Education in the AI Era

5.1 Establishing Dual Core Educational Goals

Traditional knowledge-oriented instructional logic fails to adapt to AI-supported learning environments. This review proposes taking critical thinking and problem-solving competency as two core educational objectives of STEM education. Disciplinary knowledge points are no longer the terminal standard for classroom assessment, but merely carriers for training students’ comprehensive cognitive abilities. In all inquiry tasks, complete thinking processes take precedence over standardized answers, cultivating interdisciplinary compound talents equipped with critical judgment and innovative capacity.

5.2 Multiple Support for Repositioning

This readjustment of STEM educational positioning receives three layers of theoretical, epochal and empirical support. Theoretically, social constructivism, inquiry-based learning and human-machine collaborative learning theories all advocate reducing mechanical knowledge indoctrination and strengthening high-order competency training. From an epochal perspective, basic memory and computational tasks can be fully automated by AI, making independent critical thinking humanity’s irreplaceable competitive edge. Empirically, multi-national classroom intervention experiments prove that standardized human-machine collaborative classrooms can simultaneously improve students’ two core high-order capabilities [6].

5.3 Comprehensive Transformation and Development Direction

Following the establishment of new dual-core educational objectives, synchronous adjustments are required for classroom delivery and assessment systems. Traditional classrooms revolve around knowledge mastery and exam scores; after transformation, instructional focus shifts to students’ growth in thinking capacity. One-way teacher lecturing is replaced with a tripartite model of teacher guidance, AI auxiliary support and independent student exploration. Single final paper exam scoring is substituted with a multi-dimensional process-oriented comprehensive evaluation.

6. Implementation Path of Empowering STEM Education with Generative AI

6.1 Building an AI-Supported Classroom Teaching Model

AI scaffolding teaching is constructed based on Vygotsky's zone of proximal development theory to redesign full classroom procedures. He clarified standardized AI usage norms: AI shall only provide layered inspirational prompts and is prohibited from outputting complete step-by-step problem-solving procedures [9]. When students encounter learning obstacles, AI guides them to sort out logical thinking step by step, and helps them identify logical loopholes in wrong answers after class. Long-term training under such rules can steadily advance students' deductive reasoning and critical thinking skills.

6.2 Building AI-Empowered Interdisciplinary Project-Based Learning

Chen and Wang proposed a clear labor division scheme for human-machine collaborative PBL: AI undertakes basic tasks including data collection and model simulation, while student groups independently complete higher-order links such as problem definition, comparative analysis of multiple schemes and iterative optimization of results [6]. Teachers supervise the whole process for students to verify, revise and optimize all AI-generated content, embedding training of the two core high-order competencies into every session of project learning.

6.3 Optimize the Comprehensive Evaluation System for the Entire Process

Schools can revise existing assessment rules, lower the weighting ratio of written exam scores, and place greater emphasis on evaluating students' performance in problem analysis, logical reflection, innovative scheme design and rational standardized usage of AI. Leveraging AI to record full-process classroom participation and homework revision records, combined with teacher scoring, group peer evaluation and student self-reflection summaries, multi-dimensional comprehensive grading can objectively reflect students' development of high-order core competencies.

6.4 Establish a Risk Prevention and Control Mechanism for the Use of AI Classrooms

Wu summarized a three-layer parallel risk control path: "First, formulate differentiated AI usage standards matching various classroom scenarios. Second, add result defense sessions to regular assessment links. Third, perfect campus academic integrity punishment regulations" [7]. Multiple coordinated control measures jointly mitigate risks including superficial pseudo-learning and AI-assisted academic writing misconduct, guaranteeing smooth operation of standardized human-machine collaborative STEM classrooms.

7. Conclusion

First, generative AI brings both opportunities and challenges to STEM classroom teaching; the traditional knowledge-imparting and mechanical problem-solving teaching mode cannot adapt to the current AI-integrated educational environment. Second, prominent paradigm gaps exist between domestic and international research communities: most domestic papers focus on intelligent tool development, while high-quality top-level theoretical research targeting high-order competency cultivation remains scarce. Third, STEM education requires comprehensive systemic transformation centered on critical thinking and problem-solving abilities, with synchronous upgrades to classroom teaching and evaluation mechanisms. Fourth, a completely new educational framework can be implemented relying on four supporting measures: AI scaffolding classroom design, human-machine collaborative project-based learning, full-process comprehensive assessment and AI usage risk management. Domestic teaching researchers and frontline teachers must update educational concepts and stop overemphasizing simple AI tool application alone. This review is merely a literature review, restricted by the research conditions of an undergraduate researcher. No field classroom investigations or controlled teaching experiments were conducted in primary, secondary or university settings. The proposed educational repositioning framework and four implementation paths lack supporting field data reflecting adaptive classroom effects across different regions and educational stages; follow-up empirical research is required to verify their practical operability. In addition, the literature retrieval scope of this study is limited, excluding many cutting-edge papers from niche sub-disciplines. Partial conclusions drawn herein carry one-sidedness. Future relevant research can expand literature screening

coverage and supplement targeted discussions on segmented sub-fields of STEM education. Subsequent research can advance along three major directions: First, carry out localized controlled classroom experiments to test the actual instructional effectiveness of the theoretical framework and implementation strategies proposed in this review. Second, design layered differentiated AI classroom usage schemes corresponding to distinct cognitive characteristics of primary, secondary and university students. Third, develop quantitative measurement evaluation criteria targeting the two core high-order competencies, build an integrated “Cultivation-Evaluation” research system, and continuously promote the optimization and upgrading of STEM education under the generative AI background.

References

- [1] OECD. (2022). PISA 2021 creative thinking assessment report. OECD Publishing.
- [2] Yang, M. Q. (2026). Research on the practical dilemmas of localized STEM education in China. *Journal of Southwest University (Educational Science Edition)*, (02), 45–52.
- [3] OpenAI. (2023). GPT-4 technical report.
- [4] Zimmerman, B. J. (2002). Self-regulated learning: An overview. *Theory Into Practice*, 41(2), 64–70.
- [5] Moher, D., & Liberati, A. (2009). Preferred reporting items for systematic reviews and meta-analyses (PRISMA). *PLOS Medicine*, 6(7), e1000097.
- [6] Chen, L., & Wang, Y. (2024). A literature review on artificial intelligence empowered project-based STEM learning. *E-education Research*, 43(1), 45–62.
- [7] Wu, D. J., & Liang, J. C. (2023). A review on the application, challenges and countermeasures of large language models in higher education. *Journal of Computer Assisted Learning*, 39(6), 930.
- [8] Fang, H. G., & Kong, X. M. (2022). Research on human-machine collaborative education theory in the era of artificial intelligence. *Modern Educational Technology*, (07), 36–42.
- [9] He, K. K. (2017). Connotation, types and classroom application of instructional scaffolding. *China Educational Technology*, (04), 1–9.
- [10] Zheng, Y. H., & Zhang, D. B. (2025). Bibliometric analysis of domestic literature on generative AI empowering STEM education. *China Educational Technology*, (08), 92–99.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).