

基于生成式人工智能的教育智能体如何影响编程学习效果——系统性文献综述

How Generative AI-Based Educational Agents Influence Programming Learning Outcomes: A Systematic Literature Review

向佳慧¹, 周艳^{1*}

浙江师范大学教育学院

*y.zhou@zjnu.edu.cn

【摘要】 基于生成式人工智能（GenAI）的教育智能体具有个性化反馈、知识追踪及激发动机等特点，可解决在编程学习中个性化学习及抽象概念理解困难等问题。通过系统性文献综述，本研究梳理了 31 篇相关文献，探究基于 GenAI 的教育智能体在编程学习中的对学习者的认知和非认知层面分别产生何种影响。结果显示，在认知层面，它通过代码生成、即时反馈提升了学习者的编程技能与计算思维；在非认知层面，它增强了学习动机与自我效能感。然而，结果也揭示了学段覆盖不均、学习者的学习效果会受到认知水平影响等问题。

【关键词】 生成式人工智能；编程能力；编程学习

Abstract: Educational agents based on generative artificial intelligence (GenAI) feature personalized feedback, knowledge tracking, and motivation-boosting mechanisms, which can address challenges such as personalized learning and difficulties in understanding abstract concepts in programming education. Through a systematic literature review, this study analyzed 31 relevant studies to investigate the cognitive and non-cognitive impacts of GenAI-based educational agents on learners in programming education. The results indicate that, at the cognitive level, these agents enhance learners' programming skills and computational thinking through code generation and real-time feedback; at the non-cognitive level, they boost learning motivation and self-efficacy. However, the findings also reveal issues such as uneven coverage across educational levels and the fact that learners' outcomes are influenced by their cognitive abilities.

Keywords: Generative Artificial Intelligence, Programming Competence, Programming Learning

1. 引言

在高速高质量发展的数智时代，GenAI 的出现及发展带来了前所为有的变革。传统的编程学习方法难以满足学生的个性化学习需求。编程学习所带来的挫败感可能会导致学习者学习动机的下降（Atiq & Loui, 2022），极大地限制了学生学习编程的效率以及计算思维的发展。传统的 ITS 因为定义的教学材料和基于规则的算法因素导致缺乏必要的灵活性，无法满足在学习者的多样化、动态以及对话的需求。与传统的 ITS 相比，基于 GenAI 的教育智能体具备个性化资源推荐，使其可以提供学生个性化的学习体验（Adam et al, 2025）。由于目前基于 GenAI 的教育智能体成效根据学习者的特征等情况呈现差异化结论，亟需系统化的整理。因此，本研究采用系统性文献综述法对基于 GenAI 的教育智能体如何影响编程学习效果进行梳理，找出其存在的挑战以及未来的发展方向，为后续的研究提供理论借鉴。因此，本研究提出了以下研究问题：问题一：基于 GenAI 的教育智能体对编程学习者在认知方面产生了什么

效果？问题二：基于 GenAI 的教育智能体对编程学习者在非认知方面产生了什么效果？

2. 研究方法

2.1. 文献检索、筛选及纳入

在文献检索方面，本研究以 web of science, Proquest, Science Direct 这三个数据库为主。客观并且有根据地选择 GenAI 的教育智能体在编程学习方面的研究。

考虑到 GenAI 这一技术自 2022 年出现，因此将检索时间设为 2022-2025 年。在检索词方面，以“Generative AI”, “LLMs”, “Chat GPT”, “Programming Learning”, “Programming Education”为检索词。纳入标准：(1) 明确基于 GenAI 教育智能体的研究；(2) 具有编程教育情境和学习者参与的研究；(3) 全文可获取的文献；(4) 实证研究。排除标准：(1) 非编程教育情境的文献；(2) 缺乏对编程学习者认知和非认知学习效果研究；(3) 无法获取全文；(4) 非实证研究。对所纳入的文献进行 JBI 质量评估，有 25 (80.6%) 篇处于较高水平。筛选流程见图 1。

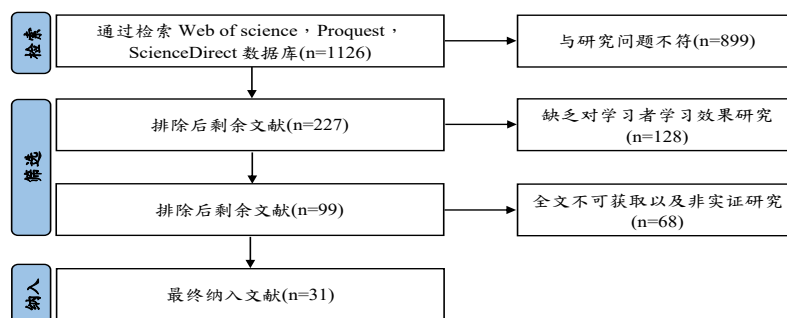


图 1 文献筛选及纳入流程

3. 研究结果

2023 年发表 2 篇 (6.5%)，2024 年发表 7 篇 (22.6%)，但有 22 篇文献 (70.9%) 发表在 2025 年，并且近五年相关文献的数量逐年增加。25 篇文献 (80.6%) 的研究学段是高等教育阶段，高中仅 3 篇 (9.7%)，初中 2 (6.5%) 篇，小学 1 (3.2%) 篇。说明基于 GenAI 的智能体在编程学习方面的研究集中在高等阶段，而缺乏 K-12 阶段的实证研究。

3.1. 问题一：基于 GenAI 的教育智能体对编程学习者在认知方面产生了什么效果？

在认知方面，基于 GenAI 的教育智能体在编程学习中可以有效的促进学生的计算思维 (N=11)。但在影响维度上存在着差异。Vishnu 等人 (2025) 认为，基于 GenAI 的教育智能体提供代码片段节省了学生的时间，从而投入更多至其计算思维等高阶认知活动中。此外，运用已有经验和知识对 GenAI 生成的内容进行批判性评估或结合思维导图可以促进学生的批判性思维和独立思考的能力 (Ye et al., 2025)。然而，Li 等人 (2025) 认为其影响有限，且在不受监管的学习环境中过度使用 GenAI 可能导致学生产生过度依赖，引发负面认知卸载，反而阻碍了学生对知识的掌握及批判性思维的发展。

在编程能力 (N=9) 和学业成绩 (N=12) 方面，基于 GenAI 的教育智能体可以提升编程中的代码的相关能力以及学业成绩。但效果会因学习者的知识水平而有所差异。通过基于 GenAI 的教育智能体生成的个性化反思报告可以提升学生的全过程编程能力 (Lee et al., 2025)。此外，Tang 等人 (2025) 提到，教育智能体通过提供动态脚手架帮助学生完善代码和加深理解来支持迭代学习。然而，ChatGPT 帮助学生学习语法和语义，而非语用学，这损害了结对编程原则，从而限制了同伴之间的交流协作及技能发展 (Groothuijsen et al., 2024)。

在学业成绩方面，大多数研究认为引入基于 GenAI 的教育智能体后，学生在概念知识、知

识收获以及学习成绩要高于其他辅助学习工具或无辅助学习的学生 (Vishnu et al., 2025)。此外, 利用智能体的提示功能或结合支架式学习可以提高学习者对概念的理解以及学习结果 (Nathaniel et al., 2025)。然而, Suh 等人 (2025) 揭示了学习者的现有水平会影响基于 GenAI 的教育智能体在编程学习中的应用效果。对于中级学习者来说, 定期运用 ChatGPT 的高质量示例可以提升其代码原创性与成绩。对于初学者来说, 运用 ChatGPT 只是简单地复制未学习的内容, 并没有提高学生的编程技能。

3.2. 问题二: 基于 GenAI 的教育智能体对编程学习者在非认知方面产生了什么效果?

基于 GenAI 的教育智能体在提升学习动机 (N=11) 与自我效能感 (N=10) 方面总体呈现积极影响。相较于传统学习模式, 智能体营造的个性化、低焦虑的环境提高了学生的积极性以及参与度 (Tang et al., 2025)。个别研究者观察到 GenAI 支持组的调试频率较低, 表明智能体并不一定能提高学生的参与度 (Li et al., 2025)。此外, 初学者认为它能增强自信心, 但会产生对未来职业产生负面影响, 而中级用户则会对未来职业持乐观态度 (Suh et al., 2025)。

自我效能感代表一个人对完成某项任务的内在信心, 它是影响编程学习成功的重要因素。Tang 等人 (2025) 认为, 基于 GenAI 的教育智能体是持续性的学习伴侣。当学生有明确的目标或者面对具有挑战性的任务时, 基于 GenAI 的教育智能体提供的个性化、实时的反馈可以增强学生的编程自我效能感, 包括信心和学习体验 (Ye et al., 2025 ; Lee et al., 2025)。

4. 讨论

本研究通过对 31 篇符合标准的实证文献分析发现, 大部分实证研究集中在高等教育阶段, 在 K-12 阶段的研究较少。一方面是因为高等教育的学生对技术类应用接触的比较多, 具备成熟的认知能力、学习周期长。另一方面是因为皮亚杰的认知发展阶段理论。低学龄阶段, 学生思维以具体运算为主, 对抽象的代码符号理解较为困难, 容易引发认知过载。学生进入中学阶段后, 学生思维进入形式思维阶段, 具备初步的推理和假设演绎能力, 但学生的自我监控能力较弱。面对基于 GenAI 的教育智能体的及时反馈, 容易进入复制粘贴思维。因此, 基于 GenAI 的教育智能体进行编程学习时, 在低年龄阶段可以结合可视化, 多样式的支架, 例如在基于 GenAI 的教育智能体中加入语音、手势动作或 AR 界面。在中学阶段可以结合提示词, 逐步引导学生进入独立解决问题的能力。

尽管 GenAI 在编程教育中展现出积极作用, 尤其是在促进计算思维、提升编程技能、增强学习动机以及提高自我效能感方面, 但其效果会受到学习者的知识水平以及使用方式等因素的影响而产生差异。对于有编程基础的学习者来说, 运用基于 GenAI 的教育智能体所提供的代码时可以初步判断代码是否有语法错误, 并将自己的编程逻辑与基于 GenAI 的教育智能体进行交互以提高编程能力。而对于初学者来说, 对智能体所提供的内容可能无法判断正误, 从而直接运用代码并缺乏与智能体进行有意义的交互。因此, 对初学者可以加入具有情景化和具身化的脚手架。比如, 将抽象的编程概念变为图像, 将重要代码以填空的形式进行引导。对于高先验知识者来说, 可以加入具有冲突的脚手架, 将有问题的编程知识呈现出来, 引导学生进行批判性评估。

此外, 若学生对基于 GenAI 的教育智能体的过度依赖可能会阻碍学生的计算思维等高阶思维能力的发展。当学生进行到代码调试时, 其编程思维以及认知会有所提高。然而, 若学生利用基于 GenAI 的教育智能体来长时期解决这些问题, 可能会丧失批判性评估以及独立思考的能力。因此, 可以在 GenAI 的教育智能体给出反馈前加入自我解释环节, 引导学生进行批判性思考。

5. 结论

基于 GenAI 的教育智能体在编程教育中的成效并非技术本身所能决定，而是依赖于教学策略、学习者特征等因素。其对学生认知层面的提升是需要根据学习者当前的认知水平进行教学设计，否则学生可能因过度依赖等因素而产生负面效果。在与基于 GenAI 的教育智能体进行交互时，能有效提升学生的学习动机、兴趣和自我效能感的非认知层面。但目前多数研究是在高等教育阶段进行。后续研究建议更多地考虑在 K-12 阶段进行研究，从而得出更为全面的结论。

参考文献

- Adam, A. B. M., Diallo, E. M., Muthanna, M. S. A., Alkanhel, R. I., Muthanna, A., & Hammoudeh, M. (2025). Generative AI-driven reinforcement learning for beamforming and scheduling in multi-cell MIMO-NOMA systems. *Physical Communication*, 72, 102771. <https://doi.org/10.1016/j.phycom.2025.102771>
- Atiq, Z., & Loui, M. C. (2022). A qualitative study of emotions experienced by first-year engineering students during programming tasks. *ACM Transactions on Computing Education*, 22(3), 1–26. <https://doi.org/10.1145/3507696>
- Groothuijsen, S., van den Beemt, A., Remmers, J. C., & van Meeuwen, L. W. (2024). AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning. *Computers and Education: Artificial Intelligence*, 7, 100290. <https://doi.org/10.1016/j.caeai.2024.100290>.
- Lee, H. Y., Chen, P. H., Lin, C. J., Huang, Y. M., & Wu, T. T. (2025). Leveraging ChatGPT for personalized reflective learning in programming education: effects on self-efficacy, higher-order thinking, and project implementation skills. *Education and Information Technologies*, 30(17), 24815-24854. <https://doi.org/10.1007/s10639-025-13733-z>
- Li, S., Liu, J., & Dong, Q. (2025). Generative artificial intelligence-supported programming education: Effects on learning performance, self-efficacy and processes. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.9932>
- Nathaniel, J., Oyelere, S. S., Suhonen, J., & Tedre, M. (2025). Investigating the impact of Generative AI integration on the Sustenance of Higher-Order Thinking Skills and understanding of programming logic in Programming Education. *Computers and Education: Artificial Intelligence*, 100460. <https://doi.org/10.1016/j.caeai.2025.100460>
- Suh, J., Lee, K., & Lee, J. (2025). Programming education with ChatGPT: outcomes for beginners and intermediate students. *Education and Information Technologies*, 1-26. <https://doi.org/10.1007/s10639-025-13542-4>
- Tang, B., Liang, J., Hu, W., & Luo, H. (2025). Enhancing programming performance, learning interest, and self-efficacy: the role of large language models in middle school education. *Systems*, 13(7), 555. <https://doi.org/10.3390/systems13070555>
- Vishnu, S., Sahil, and Garg, N. (2025), Unveiling the Role of GPT-4 in Solving LeetCode Programming Problems. *Computer Applications in Engineering Education*, 33: e22815. <https://doi.org/10.1002/cae.22815>
- Ye, X., Zhang, W., Zhou, Y., Li, X., & Zhou, Q. (2025). Improving students' programming

performance: an integrated mind mapping and generative AI chatbot learning approach. *Humanities and Social Sciences Communications*, 12(1), 1-13.
<https://doi.org/10.1057/s41599-025-04846-4>