Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Effects of Gen AI in Enhancing Pre-service Teachers' Instructional Design Competency: A Quasi-Experimental Study

Le Zhang[a], Hongchao Peng[b*] & Mingrui Zhang[a]
[a]*Department of Educational Administration, East China Normal University, China*
[b]*School of Open Learning and Education, East China Normal University, China*
*hcpeng@dec.ecnu.edu.cn

Abstract: This study investigates the impact of Generative Artificial Intelligence (GenAI) on pre-service teachers' instructional design skills, distinguishing between immediate task performance and the development of lasting competency. Employing a quasi-experimental design, this study tracked 30 pre-service teachers from a university in China over a four-week period. Data on their instructional design tasks were collected via expert ratings and self-assessments at three stages: a pre-test, a post-test one week later with GenAI-supported intervention, and a delayed post-test one week later without the tool. The results revealed three key findings: (1) The GenAI intervention significantly enhanced participants' immediate instructional design performance.(2) These gains were sustained in the delayed post-test, suggesting a genuine development of underlying instructional design competency.(3) This positive effect, however, was not mediated by the participants' AI teaching self-efficacy. These findings provide crucial empirical evidence that GenAI can be an effective tool for fostering durable professional skills in teacher education, beyond simply aiding in-the-moment task completion.

Keywords: Gen AI, instructional design competency, pre-service teachers, effects of Gen AI

1. Introduction

Generative Artificial Intelligence (GenAI) represents a paradigm shift in computing, with many strong claims about its revolutionary impact on education (Chiu, 2024; Hodges & Kirschner, 2024). A primary area of interest is its application in instructional design, where research suggests GenAI can act as a powerful scaffold, particularly for novice teachers (Berg & du Plessis, 2023). It can support the entire design lifecycle, encompassing key stages such as learner analysis, objective setting, activity and material design, and instructional evaluation (Sailer et al., 2023; Hilliger et al., 2022).

However, a critical tension exists between this potential and the practical impact on teacher development. This is due to significant challenges, including the unreliability of AI outputs and the risk of fostering over-reliance, which may diminish users' critical thinking (Ruiz-Rojas, 2023; Chan & Hu, 2023). This issue is particularly salient for pre-service teachers and raises a crucial distinction: while GenAI may enhance immediate task performance, it is uncertain whether it fosters genuine professional competency (Hodges & Kirschner, 2024). A recent study highlighted this gap, finding performance differences between groups using GenAI but failing to address the underlying development of competency (Krushinskaia, 2024). Therefore, a key unresolved question is whether GenAI acts merely as a temporary support or as a catalyst for durable skill development. To explore why this distinction might exist, we turn to self-efficacy theory (Bandura, 1997). In education, teacher self-efficacy is a known predictor of instructional quality (Zee & Koomen, 2016), and recent studies have linked AI-specific self-efficacy to teachers' willingness to adopt and persist with AI tools (Yao & Wang, 2024). This suggests AI teaching self-efficacy could be a key mediating factor in this process.

Accordingly, this research addresses three core questions: Does the use of GenAI improve pre-service teachers' instructional design performance? If so, does this improvement persist after the technology is removed, indicating a genuine enhancement of instructional design competency? Does AI teaching self-efficacy mediate the impact of GenAI on the development of this competency?

2. Research Methodology

2.1 Participants and procedures

Participants were 30 pre-service chemistry education majors (aged 19–21) from a university in Shanghai, China. This homogenous cohort was selected because all had previously completed the same "Instructional Design and Practice in Information Technology" course, ensuring baseline familiarity with instructional design principles and AI concepts. Participants were randomly assigned to an experimental (n=15) or a control group (n=15).

2.2 Quasi-Experimental Research Design and Procedures

The research employed a three-phase, longitudinal quasi-experimental design (pre-test, intervention, delayed post-test) conducted over three weeks. The independent variable was the availability of GenAI support, while the primary dependent variable was instructional design competency, measured by expert ratings and self-assessments of participants' lesson plans. AI teaching self-efficacy was examined as a potential mediator.

The procedure unfolded in three distinct phases:

Week 1 (Pre-Test): Both groups completed an initial instructional design task without any GenAI support to establish baseline competency.

Week 2 (Intervention): Following a one-week interval, both groups received a brief, conceptual training on GenAI. This was done to ensure equal awareness and mitigate motivational biases (e.g., perceived resource inequality), thereby isolating the intervention's effect to the application of GenAI, not mere knowledge of it. The control group received no hands-on practice. Immediately after this training, all participants completed the AI teaching self-efficacy scale. Subsequently, the experimental group completed the second design task with GenAI support, while the control group completed it without.

Week 3 (Delayed Post-Test): After another one-week interval, both groups completed a final design task, again without access to the GenAI tool. This phase aimed to measure the retention of skills and assess whether competency gains persisted after the technological scaffold was removed.

2.3 Experimental Setup and Data Collection

We assessed pre-service teachers' instructional design competency through three inquiry-based tasks centered on core high school chemistry topics: "Detecting Iron in Food," "Measuring the pH of Rainwater," and "Understanding Automotive Energy Sources." The content and difficulty of these tasks were validated by an expert panel using the Delphi method. The tasks were intentionally designed to be challenging, creating a genuine need for GenAI support. The selected GenAI tool was Eduaide.AI (ai.eduku.cn), chosen for its specialization in instructional design and ease of access. To structure the design process, participants were provided a template based on the 5E Inquiry-Based Instructional model (Engage, Explore, Explain, Elaborate, and Evaluate), which included relevant tips and guidelines.

2.4 Measurement and Measuring tools

To evaluate instructional design competency, we utilized the 5E Inquiry-Based Instructional Design Evaluation Scale (5E ILP), which was developed by Goldston et al (2013). This validated instrument, recommended for use with the U.S. Next Generation Science Standards (Goldston et al., 2010) , assesses inquiry-based lesson plans across five dimensions: Engage, Explore, Explain, Elaborate, and Evaluate. Each dimension is rated on a 5-point scale (1 = "very poor" to 5 = "excellent"). The scale demonstrated good reliability in this study, with a Cronbach's alpha of 0.77.

AI teaching self-efficacy was measured using a scale developed by Chou et al. (2024). This instrument assesses teachers' self-efficacy in AI-supported teaching through 25 items across five dimensions: Self-Affirmation, Teaching Enthusiasm, Diligence and Perseverance, Negative Awareness, and Positive Beliefs. Items were rated on a 5-point Likert scale (1 = "Strongly Disagree" to 5 = "Strongly Agree"). The scale showed good reliability, with a Cronbach's alpha of 0.73.

To ensure fidelity, both scales were translated into Chinese using a back-translation procedure (Duda et al., 1998).

3. Data Analysis and Results

3.1 Comparison of Pre-service Teachers' Baseline Instructional Design Competency

An independent samples t-test on pre-test scores confirmed that the experimental and control groups were equivalent at baseline. There were no significant differences in their instructional design competency, as measured by either expert ratings ($t=1.60$, $p=.12$) or self-assessments ($t=0.28$, $p=.39$). This established a valid foundation for comparing the effects of the post-intervention outcomes.

3.2 Exploring the Effect of Gen AI on Pre-service Teachers' Instructional Design

An independent samples t-test was conducted on the post-intervention scores to assess the impact of the GenAI intervention. The results revealed that the experimental group scored significantly higher than the control group on both expert ratings ($M=76.41$ vs. $M=63.28$, $p= .008$) and self-assessments ($M=67.87$ vs. $M=56.40$, $p=.007$). Furthermore, the magnitude of these differences was substantial. Large effect sizes were observed for both expert ratings (Cohen's $d=1.77$) and self-assessments (Cohen's $d=2.36$), indicating that the GenAI intervention had a strong, positive effect on improving the instructional design competency of pre-service teachers.

3.3 Verification of Substantive Enhancement in Instructional Design Competency

To assess the retention of skills, a delayed post-test was conducted one week later, during which neither group had access to GenAI. An independent samples t-test revealed that the experimental group continued to significantly outperform the control group on both expert ratings ($M=74.63$ vs. $M=66.26$, $p=.008$) and self-assessments ($M=64.00$ vs. $M=55.47$, $p=.016$). This lasting effect was substantial, with large effect sizes observed for both expert ratings (Cohen's $d=1.07$) and self-assessments (Cohen's $d = 0.93$). These findings indicate that the competency gains from the GenAI intervention were durable and persisted even after the technological support was removed.

3.4 *Mediating Role of AI Teaching Self-Efficacy*

To test whether AI teaching self-efficacy mediated the relationship between the GenAI intervention and instructional design competency, we conducted a mediation analysis using Hayes' PROCESS Model 4 with 2000 bootstrap resamples.

Table 1. Mediating Effect of AI Teaching Self-Efficacy

Path	c Total Effect	a	b	a*b (Indirect Effect)	Boot SE (Indirect Effect)	95% Boot CI (Indirect Effect)	c' (Direct Effect)
Gen AI (X) → AI Teaching Self-Efficacy (M) → Expert Rating (Y)	13.12**	12.67**	0.24	3.04	0.11	-0.01 ~ 0.44	10.07**
Gen AI (X) → AI Teaching Self-Efficacy (M) → Self-Assessment (Y)	11.47**	12.67**	-0.01	-0.14	0.08	-0.19 ~ 0.16	11.61**

*$p<0.05$ ** $p<0.01$*

As shown in Table 1, the analysis revealed no significant indirect effect of the intervention on competency through AI teaching self-efficacy. For expert ratings, the 95% bootstrap confidence interval for the indirect effect (a*b = 3.04) was [-0.012, 0.443], which contains zero. Similarly, for self-assessments, the confidence interval for the indirect effect (a*b = -0.14) also included zero [-0.190, 0.162]. Therefore, while the GenAI intervention had a significant direct effect on improving competency (see c' path in Table 1), the hypothesis that this effect is mediated by AI teaching self-efficacy was not supported.

4. Conclusion and Discussion

4.1 GenAI's Impact on Instructional Design Performance and Competency

The findings first confirm that GenAI significantly enhances the immediate instructional design performance of pre-service teachers. This aligns with research suggesting that large language models improve task performance by serving as a powerful knowledge repository and an interactive cognitive partner (Van Den Berg & Du Plessis, 2023). Through human-AI collaboration, participants engaged in knowledge retrieval and reconstruction, effectively supporting their design process (Baidoo-Anu & Ansah, 2023).

More importantly, this study demonstrates that these gains are sustainable, persisting even after the technological support was removed. This indicates a genuine development of underlying instructional design competency, a crucial distinction from mere performance enhancement. Therefore, this research provides crucial evidence that GenAI can be framed not merely as a performance tool, but as a developmental accelerator. It facilitates a transition from novice performance to expert-like competency by simultaneously modeling expertise and demanding active metacognitive participation from the learner.

4.2 The Unexpected Role of AI Teaching Self-Efficacy

Contrary to our hypothesis, AI teaching self-efficacy did not emerge as a significant mediator in the relationship between GenAI use and competency development. One plausible explanation is that the scaffolding provided by the GenAI tool was so direct and effective that it enhanced skills without needing to be channeled through the psychological construct of self-efficacy in this short-term context (Kirschner et al., 2006). It is also possible that the influence of self-efficacy requires a longer period to manifest or was overshadowed by the novelty and strength of the technological intervention itself (Mok & Moore, 2019).

4.3 Implications for Pre-service Teacher Education

Our findings offer significant implications. Theoretically, they highlight the critical difference between technology-aided performance and technology-driven competency development.

Practically, they suggest that the pedagogical approach to integrating GenAI is paramount. Merely providing access to GenAI is insufficient; teacher education programs should focus on training pre-service teachers in prompt engineering, critical evaluation of AI outputs, and the reflective integration of AI-generated content into their pedagogical frameworks. This approach is key to fostering durable professional skills rather than dependency.

4.4 Limitations and Future Directions

Building on this study's findings, several promising avenues for future research emerge to advance the understanding of GenAI's role in teacher education.

A primary direction for future research is to establish the generalizability of these findings. This would involve conducting replication studies with larger, more diverse participant samples across various institutional and cultural contexts. Furthermore, longer-term longitudinal studies are needed to track the development of both instructional competency and AI teaching self-efficacy over an entire academic term.

Another critical area for enhancement is methodological rigor, which would allow for a clearer isolation of causal mechanisms and a deeper understanding of the learning process. Such refinements could include: Incorporating pre-intervention surveys to assess participants' GenAI familiarity. This data would allow for this variable to be controlled as a covariate or used for subgroup comparisons (e.g., novices versus experienced users). Employing methods like self-report logs, think-aloud protocols, or system usage data to account for activities occurring between formal sessions, thus providing a more granular view of skill development.Utilizing a three-group design (e.g., GenAI use vs. conceptual training only vs. no-contact control) to more effectively disentangle the effects of active tool use from mere conceptual awareness.

Pursuing these directions will be crucial for developing a nuanced, evidence-based framework for leveraging GenAI to foster genuine, durable skills in the next generation of educators.

Acknowledgements

We gratefully acknowledge the support of the Smart Education Multimodal Research Laboratory. This research was funded by the Humanities and Social Science Fund of the Ministry of Education of China (Grant No. 24YJC880105), and in part by the Shanghai University Student Innovation and Entrepreneurship Training Program (Project No. 202510269119S). This work was also supported by the Special Fund of East China Normal University for Graduate Students to Attend International Conferences.

References

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52-62. https://doi.org/10.61969/jai.1337500

Bandura, A., & Wessels, S. (1997). Self-efficacy (pp. 4-6). Cambridge: Cambridge University Press.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1), 43. https://doi.org/10.1186/s41239-023-00411-8

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

Chou, C.-M., Shen, T.-C., Shen, T.-C., & Shen, C.-H. (2024). Developing and validating an AI-supported teaching applications' self-efficacy scale. Research and Practice in Technology Enhanced Learning, 19, 035. https://doi.org/10.58459/rptel.2024.19035

Duda, J.,& Hayashi, C. T.(1998). Advances in Sport and Exercise Psychology Measuresment. Morgantown: Fitness Information Technology:471-483

Goldston, M. J., Dantzler, J., Day, J., & Webb, B. (2013). A Psychometric Approach to the Development of a 5E Lesson Plan Scoring Instrument for Inquiry-Based Teaching. Journal of Science Teacher Education, 24(3), 527–551. https://doi.org/10.1007/s10972-012-9327-7

Goldston, M. J., Day, J. B., Sundberg, C., & Dantzler, J. (2010). Psychometric analysis of a 5E learning cycle lesson plan assessment instrument. International Journal of Science and Mathematics Education, 8(4), 633–648. https://doi.org/10.1007/s10763-009-9178-7

Hilliger, I., Muñoz-Merino, P. J., De Laet, T., Ortega-Arranz, A., & Farrell, T. (Eds.). (2022). Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption: 17th European Conference on Technology Enhanced Learning, EC-TEL 2022, Toulouse, France, September 12–16, 2022, Proceedings (Vol. 13450). Springer Nature. https://doi.org/10.1007/978-3-031-16290-9

Hodges, C. B., & Kirschner, P. A. (2024). Innovation of Instructional Design and Assessment in the Age of Generative Artificial Intelligence. TechTrends, 68(1), 195–199. https://doi.org/10.1007/s11528-023-00926-x

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. Educational Psychologist, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Konak, A., & Clarke, C. (2023). Augmenting Critical Thinking Skills in Programming Education through Leveraging Chat GPT: Analysis of its Opportunities and Consequences. 2023 Fall Mid Atlantic Conference: Meeting Our Students Where They Are and Getting Them Where They Need to Be Proceedings, 45117. https://doi.org/10.18260/1-2--45117

Krushinskaia, K., Elen, J., & Raes, A. (2024). Effects of generative artificial intelligence on instructional design outcomes and the mediating role of pre-service teachers' prior knowledge of different types of instructional design tasks. In International Conference on Artificial Intelligence in Education (pp. 395-400). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64312-5_49

Lau, S., & Guo, P. (2023). From" Ban it till we understand it" to" Resistance is futile": How university programming instructors plan to adapt as more students use AI code generation and explanation tools such as ChatGPT and GitHub Copilot. In Proceedings of the 2023 ACM Conference on International Computing Education Research-Volume 1 (pp. 106-121). https://doi.org/10.1145/3568813.3600138

Mok, M. M. C., & Moore, P. J. (2019). Teachers & self-efficacy. Educational Psychology, 39(1), 1–3. https://doi.org/10.1080/01443410.2019.1567070

Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta-Llovet, J., & Gonzalez-Rodriguez, M. (2023). Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix. Sustainability, 15(15), 11524. https://doi.org/10.3390/su151511524

Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning. Learning and Instruction, 83, 101620. https://doi.org/10.1016/j.learninstruc.2022.101620

Van Den Berg, G., & Du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. Education Sciences, 13(10), 998. https://doi.org/10.3390/educsci13100998

Yao, N., & Wang, Q. (2024). Factors influencing pre-service special education teachers' intention toward AI in education: Digital literacy, teacher self-efficacy, perceived ease of use, and perceived usefulness. Heliyon, 10(14), e34894. https://doi.org/10.1016/j.heliyon.2024.e34894

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. Review of Educational research, 86(4), 981-1015