

# Exploring Teachers' Integration of Generative Artificial Intelligence in Assessment Design

Pei-Yi LIN\*

*Department of Education and Learning Technology, National Tsing Hua University, Taiwan*  
\*peiyilin@mx.nthu.edu.tw

**Abstract:** This study explores the impact of integrating generative artificial intelligence into the design of alternative assessments on teachers' epistemic fluency and cognitive skills. Using a mixed-methods approach, 14 Taiwanese teachers participated in a pre- and post-test intervention supported by qualitative inquiry. Quantitative results from paired t-tests showed significant improvements in both critical and creative thinking. Epistemic fluency showed a positive upward trend but did not reach statistical significance. Qualitative content analysis further elucidated these changes, revealing that teachers exhibited high epistemic fluency through sophisticated, multi-layered verification frameworks to counter AI-generated information. The findings suggest that while AI acts as a catalyst for idea inspiration and a rigorous tool for honing judgment, its efficacy is primarily determined by the teacher's professional agency.

**Keywords:** Generative AI, epistemic fluency, alternative assessment design, critical thinking, creative thinking, human-AI collaboration

## 1. Introduction

The rapid advancement of generative artificial intelligence (GenAI) has introduced transformative possibilities and challenges within the educational landscape. As educators are increasingly encouraged to leverage AI for pedagogical innovation, particularly in designing alternative assessments, understanding their cognitive engagement and perceptions of this technology becomes crucial. Integrating AI technologies into instruction involves concepts such as epistemic fluency—the capacity to identify, navigate, and use diverse ways of knowing, which involves understanding how different disciplines express and evaluate knowledge and possessing the cognitive flexibility to adopt the perspectives of others operating within different epistemic frameworks (Morrison, & Collins, 1995), critical thinking, and creative thinking to empower teachers to improve ideas (Hong et al., 2025)—faculties that define the essence of human-led instruction in an AI-infusing age.

While AI serves as a dynamic scaffold for instructional design, concerns regarding over-reliance and knowledge quality persist (Hodges & Kirschner, 2024). This study investigates whether direct GenAI use enhances teachers' higher-order cognitive skills and explores the qualitative nature of human-AI collaboration in the modern educational landscape. Two research questions were proposed: (1) Did the teachers' epistemic fluency, critical thinking, and creative thinking change after they used generative AI to design assessments? (2) What were the teachers' perceptions of generative AI after using it?

## 2. Methods

### 2.1 Participants and context

This study was conducted as part of an alternative assessment course for in-service teachers at the College of Education, a national university in northern Taiwan. The participants were 14 in-service teachers pursuing master's degrees, teaching in elementary and secondary schools in northern Taiwan. 71% of participants were elementary school teachers. 25% of the participants were male. There were four Chinese teachers, three English teachers, two social science teachers, two science teachers, one math teacher, one

integrative activities teacher, and one special education teacher. Their average age and teaching experience are 28.5 years ( $SD = 3.28$ ) and 5.35 years ( $SD = 3.34$ ), respectively. Regarding the frequency of using GenAI during the current semester, 50% of participants used it 1-5 times. The remaining participants were distributed across higher-frequency tiers: 14.3% reported using GenAI 6 to 10 times, another 14.3% used it 11 to 15 times, and 21.4% demonstrated high engagement by using the tools more than 15 times.

In this study, they completed pre- and post-test surveys to assess their performance in using AI to design alternative assessments in the classroom. In this course, they used generative AI to design alternative assessments twice based on their teaching practices. The researcher, acting as the course instructor, guided the participants in designing assessment rubrics based on the teaching themes first. Then, they used generative AI to design the assessment rubrics. Finally, the teachers synthesized their own designs and those using generative AI, analyzed and compiled the results, and proposed the final alternative assessment design. Before the course, they signed informed consent forms.

## 2.2 Instruments

This study employed pre- and post-surveys and open-ended questions to measure the participants' epistemic fluency, critical thinking, and creative thinking. The survey items were adapted from epistemic fluency (nine items, Cronbach's alpha = 0.90) (Lin et al., 2022), critical thinking (four items, Cronbach's alpha = 0.82), and creative thinking (four items, Cronbach's alpha = 0.88) (Chai et al., 2015). At the end of the post-survey, three open-ended questions were posed: (1) What are your experiences in using AI in teaching? (2) What are your perceptions when using AI? (3) How did you integrate AI-generated information into teaching?

## 2.3 Data analyses

The data includes quantitative and qualitative data. The quantitative data were analyzed using SPSS to compare participants' perceptions of using generative AI in assessment design. A paired t-test was used to examine changes in teachers' perceptions. Content analysis was applied to examine participants' responses to the open-ended questions.

## 3. Results

### 3.1 Teachers' epistemic fluency, critical thinking, and creative thinking towards AI

A paired t-test was performed to analyze teachers' perceptions of epistemic fluency, critical thinking, and creative thinking towards AI after they used generative AI to design alternative assessments. Based on the results, there were no significant differences in their epistemic fluency (pre-survey  $M = 5.27$ ,  $SD = 0.60$ ; post-survey  $M = 5.85$ ,  $SD = 0.98$ ;  $t = -2.01$ ,  $p = 0.06$ ). However, the effect size ( $d = -0.54$ ) suggests a moderate practical impact, and the results indicate a marginal trend toward improvement. Participants' critical thinking showed a significant improvement (pre-survey  $M = 5.89$ ,  $SD = 0.67$ ; post-survey  $M = 6.18$ ,  $SD = 0.67$ ;  $t = -2.93$ ,  $p < 0.05$ ). The effect size, as measured by Cohen's  $d$ , was  $-0.78$ , indicating a large effect. The participants' creative thinking also showed significant enhancement (pre-survey  $M = 5.43$ ,  $SD = 0.89$ ; post-survey  $M = 5.93$ ,  $SD = 1.07$ ;  $t = -2.22$ ,  $p < 0.05$ ). The effect size was medium to large ( $d = -0.59$ ). The results provide evidence that the use of GenAI to assist teachers in designing alternative assessments effectively fosters cognitive capacities.

The large effect size in critical thinking points to GenAI as an effective catalyst for developing the analytical approaches used to navigate and synthesize complex, multi-perspective information. Furthermore, the growth in creative thinking indicates an increased capacity for generating innovative pedagogical designs. Although the gains in epistemic fluency were not statistically significant, the moderate effect size suggests that participants began to bridge different ways of knowing through GenAI-supported instructional design.

### 3.2 Teachers' experiences and perceptions of using AI to design assessment

Qualitative data from open-ended responses were analyzed using an open coding strategy, yielding a framework across four core dimensions, shown in Table 1.

Table 1. Coding schemes and examples of the open-ended questions

Concept	Sub-dimension	Example
Epistemic fluency	Idea organization; teacher's role as navigator	T8: To clarify my chaotic thoughts, I ask AI to help me organize and refine my ideas.
Critical thinking	Accuracy; expertise judgment; verification process	T5: I check relevant information in search engines or books for double verification.
Creative thinking	AI as a catalyst for inspiration	T13: AI inspires me with ideas that I have not thought of before, but I have to consider whether it is practical.
Utility	Efficiency gains	T1: The use of AI is time-saving.

Qualitative findings reveal that teachers prioritize mitigating AI-driven misinformation by rigorously verifying and detecting bias. For instance, participants like T12 exhibit advanced epistemic fluency by implementing a tripartite verification framework (internal, external, and experiential), reflecting a sophisticated grasp of knowledge production. While AI serves as a valuable tool for brainstorming and resource design, the teacher's professional judgment remains the final filter, ensuring that all AI-generated content is appropriately tailored to the specific classroom context.

To conclude, the quantitative analysis reveals that AI collaboration significantly boosts critical thinking and creative thinking. While qualitative data shows that teachers act as vigilant navigators, verifying and cross-referencing AI-generated information. The results indicated that their epistemic fluency is still developing, suggesting that the ability to fluently navigate and switch frameworks between AI-generated information and human disciplinary expertise is a higher-order skill that requires more prolonged internalization than immediate fact-checking or brainstorming. Future teacher training programs should focus on shifting teachers' roles from passive users to expert evaluators, enhancing their professional agency, and recognizing their unique role as the final decision-makers.

### Acknowledgements

Thanks to the participating teachers. This study was supported by National Science and Technology Council grants in Taiwan to the Department of Education and Learning Technology at the National Tsing Hua University (NSTC 112-2410-H-007-032-MY3).

### References

- Chai, C. S., Deng, F., Tsai, P. S., Koh, J. H. L., & Tsai, C. C. (2015). Assessing multidimensional students' perceptions of twenty-first-century learning practices. *Asia Pacific Education Review*, 16(3), 389–398. <https://doi.org/10.1007/s12564-015-9379-4>
- 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>
- Hong, H. Y., Chen, M. J., Chang, C. H., Tseng, L. T., & Chai, C. S. (2025). AI-supported idea-developing discourse to foster professional agency within teacher communities for STEAM lesson design in knowledge building environment. *Computers & Education*, 229, 105241. <https://doi.org/10.1016/j.compedu.2025.105241>
- Lin, P. Y., Chai, C. S., Di, W., & Wang, X. (2022). Modeling Chinese teachers' efficacies for the teaching of integrated STEM with interdisciplinary communication and epistemic fluency. *Frontiers in Psychology*, 13, 908421. <https://doi.org/10.3389/fpsyg.2022.908421>
- Morrison, D., & Collins, A. (1995). Epistemic fluency and constructivist learning environments. *Educational Technology*, 35(5), 39–45.