

GenAI in Learning, Teaching and Assessment

17

Instructors

Dr Siyu Zhu & Professor Xinhua Zhu

Department

Language Science and Technology (LST)

Subject Area

Assessment Feedback

Why did the instructor use GenAI for learning and teaching?

Within the context of teacher education, Xinhua's team was interested in exploring how ChatGPT could be integrated into writing instruction to support pre-service teachers (PSTs) in developing more effective feedback practices.

The intervention led by his team was designed in response to well-documented challenges in this area. Although high-quality feedback is central to writing development, PSTs often experience difficulties in providing comments that are sufficiently detailed, balanced, and pedagogically effective. The team therefore investigated ChatGPT to determine whether it could address both the pedagogical and practical concerns PSTs face when grading. Crucially, the objective of the investigation was not to replace teacher judgement, but to explore whether AI-supported tools could strengthen PSTs' capacity to provide richer, more varied, and more effective writing feedback.

How was GenAI used in this scenario?

The investigation involved 30 Master of Arts students studying Chinese language education at a public research university in Hong Kong. Participants took part in a two-week intervention focused on ChatGPT-assisted writing feedback.

Intervention Design

The investigation used a pre-test/post-test design. Participants were asked to evaluate student essays and provide revision-oriented feedback with the assistance of ChatGPT. They were explicitly required to revise and refine the ChatGPT-generated feedback rather than copying it directly.

Two workshops were delivered during the intervention. These workshops were informed by the TPACK framework, integrating:

- **Technological knowledge (TK):** effective prompt-writing strategies.
- **Pedagogical knowledge (PK):** principles of effective writing feedback.
- **Content knowledge (CK):** understanding the concepts of writing feedback.

Workshop Focus

The workshops introduced participants to:

- the affordances and limitations of GenAI and ChatGPT;
- principles of effective prompt-writing;
- general feedback-giving principles;
- distinctions between lower-level and higher-level feedback;
- different feedback types, including summary, praise, criticism/problem, explanation, specific solution, and general suggestion.

A 3-Step Process

Participants were trained to use a structured three-step process:

1. **Preview** the student writing to develop an initial understanding of the essay and identify preliminary feedback priorities.
2. **Prompt and refine** ChatGPT output by providing the essay context and assessment criteria, then follow up with prompts as needed.
3. **Review and revise** the generated feedback using professional judgement to ensure accuracy, appropriateness, and pedagogical value.

This approach framed ChatGPT as a support tool within a broader process of human evaluation, refinement, and decision-making.

What was the impact on student learning?

Using an explanatory sequential mixed-methods design, the team investigated changes in the 30 PSTs' feedback levels and types before and after the intervention. The intervention produced several notable changes in their feedback practices.

Improvements in Feedback Levels

Post-test results showed significant increases in feedback related to:

- **grammar**
- **ideas and elaboration**
- **style**

There was also a marginal increase in feedback on word choice. The most substantial changes occurred in higher-level feedback, particularly regarding ideas, elaboration, and style. This is pedagogically significant because novice teachers often focus primarily on surface-level corrections rather than the deeper aspects of writing quality.

Improvements in Feedback Types

The study also found significant increases in:

- **explanation**
- **general suggestion**

These findings indicate that, following the intervention, PSTs were more likely to explain why an issue mattered and provide broader guidance to support student revision.

Combined Changes

When feedback levels and feedback types were examined together, significant increases were found in:

- **grammar + explanation**
- **word choice + general suggestion**
- **ideas and elaboration + explanation**
- **ideas and elaboration + general suggestion**
- **style + praise**
- **style + explanation**
- **style + general suggestion**

These patterns suggest that the intervention did not simply increase the amount of feedback given; rather, it improved the focus, and pedagogical orientation of that feedback.

Broader Learning Gains

Interview data indicated that participants also developed:

- **greater confidence in using ChatGPT;**
- **stronger prompt-writing skills;**
- **an improved ability to evaluate and edit AI-generated content;**
- **an increased awareness of the need for critical oversight when using GenAI.**

Overall, the investigation suggests that ChatGPT, when supported by structured pedagogical training, can help PSTs provide more substantive and reflective feedback on writing.

What were the challenges encountered during the implementation and what solutions were used?

The implementation also revealed several challenges that are highly relevant for academic staff considering GenAI integration.

Challenge 1: Reliability and accuracy of AI-generated feedback

ChatGPT did not always provide accurate, nuanced, or contextually appropriate comments. This was especially evident in relation to Chinese grammar, rhetorical features, and disciplinary judgement.

Solution:

Participants were trained to treat ChatGPT output as provisional rather than authoritative. They were required to review, filter, revise, and supplement the AI-generated feedback before using it in its final form.

Challenge 2: Limited student skill in prompting

PSTs did not automatically know how to obtain useful responses from ChatGPT. The quality of AI-generated feedback depended heavily on how prompts were framed.

Solution:

The workshops provided explicit training in prompt engineering, particularly:

- giving specific instructions;
- breaking tasks into smaller parts;
- prompting ChatGPT to check or improve its own output.

These strategies were reported by participants as among the most useful elements of the intervention.

Challenge 3: The need for human judgement in selecting and synthesising feedback

Even when ChatGPT generated potentially useful comments, PSTs needed to decide what to retain, modify, reorganise, and remove. This required both pedagogical understanding and critical reflection.

Solution:

The intervention included scaffolds such as checklists, worked examples, instructional videos, and a 3-step feedback model. These resources supported PSTs in making informed decisions and strengthened their ability to work critically with GenAI.

Challenge 4: Limited growth in specific solutions

Although explanations and general suggestions increased significantly, specific solutions did not. This may reflect both the technical limitations of ChatGPT-3.5 in Chinese and the participants' preference for less directive feedback.

Solution:

The intervention recommends further research and staff development focused on how to balance open-ended guidance, explanatory feedback, and explicit revision advice. This is particularly important in contexts where cultural or pedagogical norms may favour indirect guidance.

Challenge 5: Contextual limitations

The investigation was based on a relatively small sample, a single disciplinary area, and one writing genre; it also did not include a control group.

Solution:

The team recommends further investigation across different disciplines, writing tasks, student proficiency levels, and alternative GenAI tools. The team also notes the importance of studying how learners ultimately interpret and apply AI-supported feedback.