International Journal of TESOL Studies (2025)
251108, 1-26 https://doi.org/10.58304/ijts.251108

Article

Leveraging Hybrid Intelligence: How Teacher Agency Influences Behavioural Engagement in Human-AI Collaborative Writing Feedback Practices

Chunrao Deng
Shuyang Lin
Esther Ka-man Tong*
College of Professional and Continuing Education, The Hong Kong Polytechnic University, Hong Kong SAR, China

Jingjing Ma
The Hang Seng University of Hong Kong, Hong Kong SAR, China

Received: 1 August, 2025/Received in revised form: 1November, 2025/Accepted: 12 November, 2025/Available online: 29 November, 2025

Abstract
The prevalent use of AI technology in education has sparked versatile innovations for pedagogical enhancement. Teachers possess unique and irreplaceable qualities that are crucial to education. In the context of L2 students' genre-based writing, it remains an open question how GenAI and human teachers' feedback can complement each other to enhance students' learning outcomes. This study compares 14 feedback reports from teacher markers who utilised AI-recommended action points to varying degrees with the original AI-generated feedback report of a sample book review, along with data from a focus group interview and 12 reflective journals. It explores how teachers and Generative AI collaboratively construct feedback through an interactive AI-assisted platform tailored for genre-based writing. This study highlights the importance of teacher engagement and teacher agency in AI-enabled settings. The results reveal the complementary strengths of Generative AI (GenAI) and teachers in feedback practices. However, its effectiveness relies on teachers' critical engagement and familiarity with technology. A balanced approach is essential to maximise benefits and address challenges in AI-assisted feedback. Based on our findings, a theoretical model illustrating teachers' behavioural engagement with GenAI feedback has been derived to evaluate the effectiveness of AI-assisted feedback. This model highlights teachers' engagement with the GenAI feedback based on their agency, influenced by their capacity and willingness to perceive and shape the affordances of GenAI feedback. This research contributes to the ongoing conversation on the future of feedback delivery in an AI-driven world.

Keywords
Human-AI collaborative writing feedback, teacher agency, engagement, genre-based pedagogy, affordance

*Corresponding author. Email: esther.tong@cpce-polyu.edu.hk

1 Introduction

The rise of artificial intelligence (AI), particularly generative AI (GenAI), has prompted educators and researchers to rethink traditional feedback practices in writing instruction. In writing pedagogy, teacher feedback plays a key role in facilitating learners' academic literacy and writing proficiency development (Hyland & Hyland, 2006). Teachers are expected to actively supplement ChatGPT feedback with empathetic, content-focused guidance (Teng, 2024). However, not all teachers are able to fully harness technological affordances in their feedback practices (e.g., Li, 2021). With the increasing integration of GenAI tools in education, how teachers engage in their feedback processes within AI-assisted contexts has emerged as a pertinent issue to address.

Empirical studies have identified differences between teacher feedback and GenAI feedback (Guo & Wang, 2023). Teacher feedback has been found to be of higher quality (Steiss et al., 2024) and more conducive to successful student revisions (Zou et al., 2025). Studies that investigate the combined use of teacher and GenAI feedback (e.g., Han & Li, 2024) underscore teacher agency in shaping collaborative feedback practices. However, how teachers and AI can complement each other in providing feedback on genre-specific writing tasks remains underexplored. In addition, developing disciplinary academic writing proficiency presents a significant challenge for second language students who are studying in English-medium instruction (EMI) settings. Students are expected to master genre-specific conventions that encompass the organisational structures, language styles and lexico-grammatical features associated with respective disciplines (Hyland, 2004; Tong et al., 2023). To help students navigate genre conventions in their disciplinary communities, Content and Language Integrated Learning (CLIL) approaches have been introduced, and one approach is through feedback delivery, with language instructors supporting content teachers (Lyster & Ballinger, 2011). Previous research on AI feedback has largely focused on generic genres such as argumentative essays, leaving discipline-specific genre writing largely unexamined. Yet, providing feedback on disciplinary genres requires nuanced understanding of rhetorical conventions, genre norms, and field-specific language use – areas where teacher expertise is essential (Hyland, 2000; Nesi & Gardner, 2006; Tong et al., 2023). Our study is situated within this context, where language instructors, with the assistance of AI tools, provide feedback on students' draft assignments for discipline-specific subjects. Among the limited research on AI-teacher collaboration, teacher agency has been found to influence how teachers identify themselves and use AI (Teng & Yip, 2025). It is thus important to investigate how teachers enact their agency and engage with AI for feedback provision, an underexplored area.

This study addresses research gaps by investigating teacher engagement with GenAI feedback in their marking processes and the influence of teacher agency on their engagement within discipline-specific writing contexts in higher education. Specifically, it explores:

RQ1: To what extent do teachers engage with AI-generated feedback in their marking practices?

RQ2: How does teacher agency influence human-AI collaborative feedback decisions?

This study contributes to existing research on human-AI collaborative feedback by proposing a theoretical model that illustrates teachers' varying levels of behavioural engagement with GenAI feedback, shaped by teacher agency, which in turn is mediated by multiple teacher and GenAI factors. With the discussion of strategies for leveraging hybrid intelligence in second language assessment and insights into how teacher agency shapes their engagement in human-AI collaborative writing feedback practices, this study offers implications for teacher professional development, AI tool design, and genre-based pedagogies for leveraging human-machine intelligence.

2 Literature Review

2.1 Teacher engagement with GenAI feedback

In the GenAI era where GenAI tools are transforming feedback practices, it is crucial to examine the roles of both GenAI and teacher feedback. Research has identified a multitude of benefits of GenAI feedback. For example, Teng (2024) summarised that ChatGPT's timely and customised feedback facilitates students in improving writing proficiency and confidence. Teng (2025) studied 40 EFL learners from a Macau university, with half engaged with ChatGPT during in-class writing and the other half engaged with peer-collaboration under teacher guidance for the writing tasks. It was found that ChatGPT feedback exerted significant impacts on writing motivation, self-efficacy, engagement, collaborative writing tendency and metacognitive awareness.

Studies have also compared the effectiveness of teacher feedback and ChatGPT feedback. Steiss et al. (2024) revealed that the feedback from well-trained teachers is of higher quality. Fan et al. (2024) found that students perceived teacher feedback as more reliable for genre-specific assignments, as it facilitates critical thinking and provides specific guidance based on assignment criteria. Guo and Wang (2023) compared teachers' and ChatGPT's feedback on a 300-word argumentative essay and identified differences in feedback quantity, distribution, and type. ChatGPT provided a significantly larger amount of feedback, evenly distributed across various aspects of essays, while teacher feedback mainly addressed content and language aspects; ChatGPT offered more direct revision suggestions and praise, whereas teachers provided more informative and query feedback. Zou et al. (2025) found that teacher feedback and ChatGPT feedback offered distinct advantages for student revision. In this study, 20 Chinese university students received both teacher and ChatGPT feedback on their argumentative essay drafts and revised their drafts accordingly. The analysis of their revisions revealed that students engaged more with teacher feedback and achieved a higher successful revision rate. They engaged with teacher feedback more to address language and content issues, while engaged with ChatGPT more for improving organisation. Questionnaire results indicated that students generally preferred teacher feedback for its depth and personalised nature. The study advocates integrating teacher and ChatGPT feedback to enhance the overall quality of writing feedback.

Several studies have examined the combined use of teacher and AI feedback, concluding that AI feedback supplements teacher instruction and enhances feedback efficiency (Han & Li, 2024; Han & Sari, 2024). Han and Sari (2024) compared the pre-test and post-test persuasive essay performance of two classes of Turkish university students: one receiving full teacher feedback only and the other receiving combined automated (Criterion) and teacher feedback. While no significant difference was found between the two groups in terms of overall writing quality, the group who received combined automated-teacher feedback significantly outperformed in grammar and mechanics. Student reflections highlighted their appreciation for the immediacy and accessibility of automated feedback, which may have fostered self-regulated learning and contributed to improved performance. The integration of AI-powered feedback (e.g., Criterion) with teacher feedback has been found to be more effective than teacher feedback alone for addressing grammar and mechanics (Han & Sari, 2024). Han and Li (2024) explored the effectiveness of ChatGPT-supported teacher feedback on 102 Chinese university students' argumentative and expository essays. ChatGPT was first used to identify common errors and provide holistic rhetorical feedback based on specific prompts. Four teachers then built on ChatGPT outputs to deliver their own feedback. Students' revisions displayed high engagement with this feedback type, supporting the effectiveness of the teacher plus AI model. However, the study did not address how teachers exercised their agency in adapting ChatGPT-generated feedback. These studies have demonstrated the distinct advantages of teacher feedback and GenAI feedback. While existing literature underscores the benefits of teacher-AI collaboration and the importance of teacher agency in utilising GenAI, how teachers engage with GenAI feedback remains under-researched. Moreover, previous

research has not explored teacher-AI collaborative feedback provision for discipline-specific genres, a gap that warrants attention due to the potentially higher demands on teacher agency in these contexts.

Given the limited research on teacher engagement with GenAI feedback, studies on student engagement with feedback are drawn upon to shed light on the current study. Learner engagement with feedback, defined as students' response to feedback (Ellis, 2010), has attracted increasing attention in the field of writing (Cheng et al., 2023; Huang & Teng, 2025; Zhang & Hyland, 2023). Learner engagement is a multi-faceted phenomenon, which includes the three dimensions of cognitive, affective, and behavioural engagement (Ellis, 2010). Cognitive engagement refers to students' processing of feedback and the cognitive and metacognitive operations employed to address feedback. Affective engagement refers to students' attitudes towards feedback. Behavioural engagement refers to learners' revision behaviours in response to feedback. As teachers are expected to critically engage with GenAI feedback, their approach to this feedback resembles how learners engage with external feedback. In particular, teacher participants must evaluate the quality of the GenAI feedback and decide how to use it, whether to apply, remove, add, or edit the AI-recommended action points. For the purpose of the study, teachers' cognitive engagement with AI feedback is defined as their cognitive processing of such feedback and strategies employed to evaluate its suitability. Teachers' affective engagement is defined as their attitudes towards AI feedback. Teachers' behavioural engagement is defined as their actions in response to AI feedback. This paper focuses on the behavioural dimension of teacher engagement with AI feedback, given its direct influence on feedback provision to students.

2.2 An ecological perspective on teacher agency and teacher engagement with GenAI feedback

To examine the influence of teacher agency on teacher engagement with GenAI feedback, this study adopts an ecological perspective, which emphasises the interrelationship between organisms (e.g., humans) and the surrounding environments (Gibson, 1986; van Lier, 2000). Research on English language teachers has explored the concept of agency (e.g., Teng, 2019). From an ecological perspective, agency can be defined as "the active engagement of individuals with aspects of their contexts-for-action" (Biesta & Tedder, 2007, p.132). Individuals engage with aspects of their contexts-for-action by perceiving affordances in the environment for action-taking. Here, affordances refer to perceived opportunities for action provided by the environment or functionally significant properties of the environment perceived by an actor (Gibson, 1986). To exercise agency, individuals need to have (a) the capacity to perceive possible affordances for action and (b) the intention to act on these affordances (Gibson, 1986; Reed, 1993). In other words, capacity and willingness are essential for individuals to exercise agency by perceiving the affordances in the environment and acting on them. The alignment between the agent's characteristics (i.e., capacity and willingness) and the opportunities for action in the environment is thus crucial to realising affordances which will then be acted on. In addition, individuals may exercise agency by perceiving possible opportunities for others' action taking and then shaping these possible affordances for them to use (Kyttä, 2004). From an ecological perspective (Gibson, 1986; van Lier, 2000), teachers are active agents in utilising AI for feedback provision based on their capacity and willingness. Teacher engagement with AI feedback can be seen as a process in which they perceive and utilise the learning opportunities afforded by AI feedback and, if necessary, shape new affordances when they collaborate with AI to provide feedback to students. If teachers' capacity and willingness to give helpful feedback align with the learning opportunities afforded by AI feedback, they may perceive the affordances of AI feedback and then take action (e.g., sharing useful AI feedback with students). However, given the limitations of GenAI in generating high-quality feedback on discipline-specific writing (Fan et al., 2024), it is also possible that teachers may perceive limited or even no affordances in AI feedback. In this case, it is likely that teachers, based on their capacity and willingness, may perceive

and create possible learning affordances in their own feedback, which may supplement AI feedback, to facilitate students' learning. Teacher agency and teacher engagement with AI can also be explained by distributed agency, which sees agency as distributed across human, technologies, and institutions (Jones, 2022; Godwin-Jones, 2024). Working with the assistance of GenAI tools is seen as a collaborative product with multiple agents – human and digital tools – negotiating meaning and agency (Godwin-Jones, 2024). How agency is distributed across parties depends on how users engage with GenAI tools (Tsao & Nogues, 2024). Important factors that affect teacher capacity in feedback provision and the distributed agency between teacher and AI include teacher feedback literacy and critical digital literacy. These two factors will be reviewed below.

2.3 Teacher feedback literacy

Teacher feedback literacy is an important capacity-related factor that mediates teacher agency in the context of AI-assisted feedback practices. It refers to the knowledge, skills, and dispositions that enable teachers to provide quality feedback and support students' uptake of that feedback (Lee, 2019). In L2 writing classrooms, teacher feedback literacy encompasses teacher values, attitudes, and goals concerning feedback as well as what they need to do before, during, after and beyond feedback. Particularly relevant to the study is teachers' ability to prioritise higher-order communication goals and to use appropriate feedback techniques during feedback processes. For example, teachers need to draw upon their knowledge of high-quality feedback to provide actionable comments that are concrete and text-specific (Lee, 2021). Technology plays a significant role in shaping teacher feedback literacy, particularly in the context of L2 writing instruction (Carless & Winstone, 2020). Feedback literate teachers can leverage technology in the three dimensions of teacher feedback literacy: create conditions that support feedback uptake, enhance dialogic exchanges, and manage pragmatic compromises (Carless & Winstone, 2020). The emergence of technological tools has brought opportunities for teachers to enhance feedback practices (Cunningham & Link, 2021) or use their potential to facilitate students' engagement with and learning from feedback; however, it was found that not all teachers are able to fully harness technological affordances (Jiang et al., 2020; Li, 2021). In Automatic Writing Evaluation (AWE) contexts, teachers must mediate its use through scaffolding and complementary feedback to maximise learning potential (Jiang et al., 2020). In teacher-AI collaborative writing feedback practices, it is likely that teacher feedback literacy enables teachers to exercise agency to engage with AI feedback so that students can improve their writing.

2.4 Critical Digital Literacy (CDL)

Critical Digital Literacy (CDL) constitutes another important capacity-related factor that shapes teacher agency in utilising AI and AI-generated feedback. CDL refers to the ability to critically engage with digital tools, recognising their embedded ideologies, biases, and power structures (Darvin, 2025). CDL views AI use as a form of literacy and a socially situated practice, requiring users to understand how AI tools function and the implications of their algorithms and data sources (Darvin, 2025). Without CDL, users (e.g., learners) may be disempowered by GenAI when they delegate tasks like drafting or summarising entirely to GenAI (Moorhouse et al., 2025), resulting in reduced engagement with the learning process. False empowerment occurs when users feel empowered by GenAI but actually lack critical digital literacy to understand its limitations and biases (Moorhouse et al., 2025). While CDL has been found to influence how L2 writers engage with GenAI (Moorhouse et al., 2025), it is also relevant to the discussion of teacher agency in the context of teacher-AI collaborative writing practices. For instance, Zhang and Dikilitaş' (2025) case studies of two novice teachers show that one teacher saw GenAI literacy as peripheral, while the other actively engaged in GenAI training and peer networks to

develop stronger critical GenAI literacy and agency. Teachers' varying levels of critical AI literacy are likely to influence their agency in engaging with GenAI feedback.

3 Methodology

3.1 Context and participants

This study was conducted in the academic year of 2024-25 at a self-financing tertiary education institution in Hong Kong. Fourteen language instructors of a writing intensive subject titled "Fiction and Life: Understanding Human Development" participated in this study in which 64.3% of them were female (N=9) and 35.7% were male (N=5). In this course, 662 students were required to submit two drafts of an academic book review, one of approximately 700 words and the other around 1500 words, along with a final paper of about 1800 words. The language instructors used a customised AI-assisted feedback platform to provide a feedback report with 6 to 8 actionable points on students' writing drafts completed to fulfill the English writing requirements of the bachelor's degree study. Students were expected to revise their writing in the next draft based on the teacher-AI cooperative feedback.

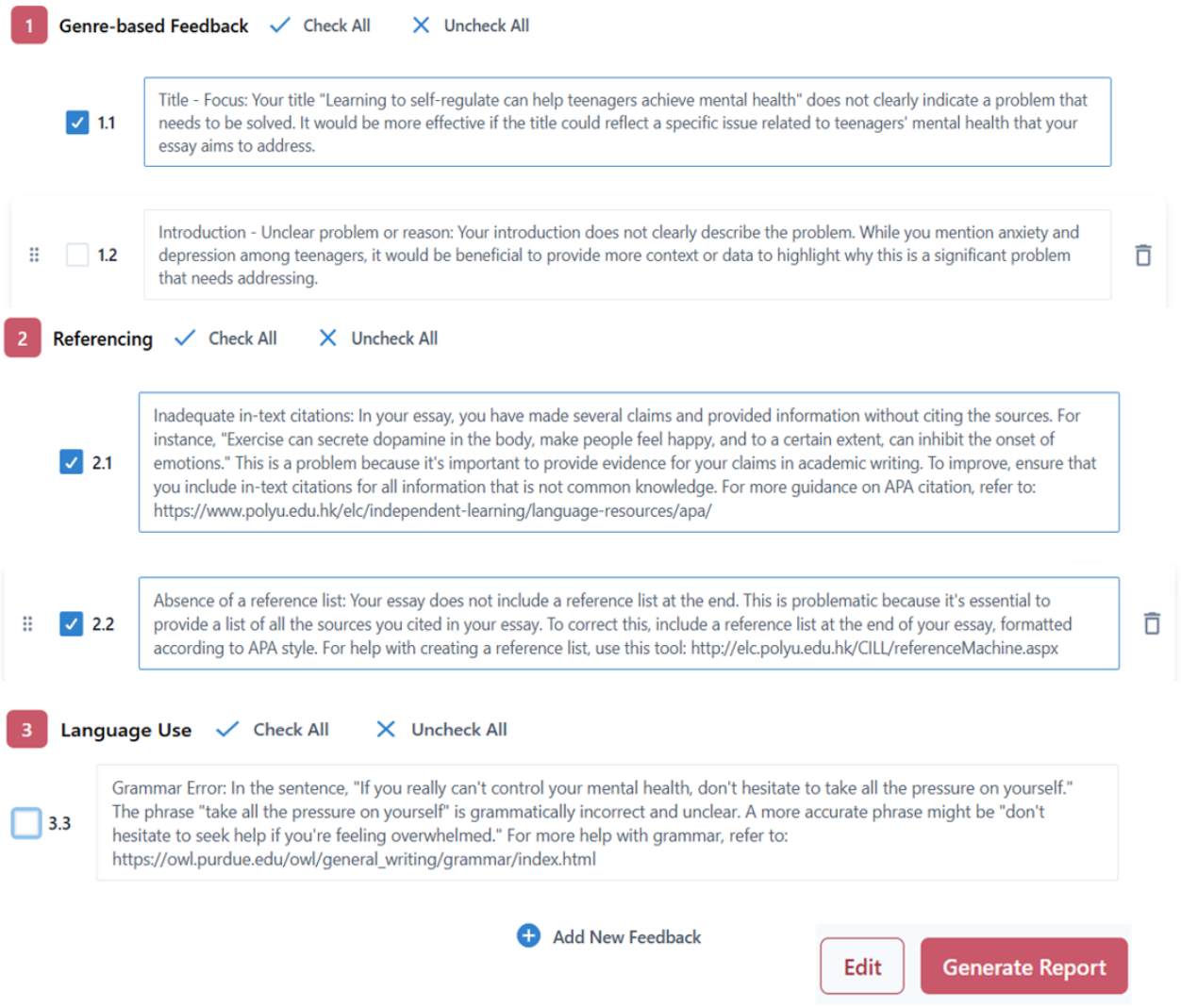
The AI-assisted academic writing platform is powered by generative artificial intelligence (GAI) technology, ChatGPT 4.0. It is developed to provide English as a second language learners with constructive assessment for learning writing support not bounded by time and location. Unlike other generative AI-enhanced proofreaders like Grammarly, this platform supports a genre-based pedagogy and provides constructive feedback on disciplinary genres that guide students' active response to writing feedback, enhancing their textual representations of disciplinary knowledge and applications through human-AI collaborative feedback. It integrates Azure cloud services and a private cloud to enhance the alignment of the discipline-specific subject assessment requirements and AI-generated feedback in a secure environment.

As shown in figure 1, the platform enables language instructors to select, edit and deselect the AI-generated feedback on three aspects, namely genre-based issues, referencing formats, and language use. Language instructors can also add and prioritise comments based on their professional judgement, pedagogical language focuses and understanding of individual students' academic backgrounds and performance. Powered by ChatGPT 4.0, the functions of this platform can facilitate the calibration and transcend the quality of feedback outputs of the language instructors. The development of the platform is funded by the government. At the time of data collection, the platform was still being constructed. Teachers' comments about their experiences using the platform were utilised for AI training.

To better prepare language instructors for using the AI-assisted writing feedback platform, all instructors were required to participate in a 2-hour training workshop. The workshop covered rationale for adopting AI, introduced platform functionality and pedagogical design, and provided step-by-step guidance on how to view, edit, select, deselect, prioritise, and add feedback points. Instructors also learned how to generate and release feedback points using the platform. After the training session, instructors were invited to review the same student book review sample. A user manual was provided to all instructors, and technical support was available to ensure they understood how to use the available functions on the platform. The dataset comprises the marking of one random sample from 662 students' draft 1 writing in week 5 of the semester. The AI-generated feedback focused on three main categories: genre-based feedback, referencing, and language use. The AI generated feedback report had approximately 500 words in length. A total of 14 action points were recommended, with four to five points in each category. Each action point was presented with a heading, a description of the area that requires improvement, and relevant resources for self-study.

Figure 1

Screenshot of the AI-Assisted Feedback Platform



3.2 Research design and procedures

To examine teacher engagement with AI-generated outputs on this platform and the role of agency, teachers' feedback reports were compared against the original AI-generated feedback report to identify patterns of teacher engagement with AI-generated feedback outputs. Data from a focus group interview and reflective journals were triangulated to assess teacher agency in an AI-supported writing assessment context. This diverse data collection provided a comprehensive understanding of the impact of AI feedback on teaching practices. Table 1 summarises the research procedures and data collection methods used in this study. Teacher feedback reports and interview data addressed Research Question 1 (RQ1), while reflective journals and interview feedback were analysed for Research Question 2 (RQ2).

Table 1 below presents the research procedures and data set of this study. The research procedure used a multi-faceted approach to gain insights into teachers' experiences with the AI feedback platform. In Week 5, we analysed 14 teacher feedback reports, comparing them with the original AI-generated feedback to assess alignment and discrepancies. The following week, we conducted a 1.5-hour focus group interview with these teachers to gather in-depth perspectives. The following questions were included in the semi-structured interview:

1. What procedures did you go through when marking the essays with the assistance of the generative AI-powered academic writing platform?
2. How did you utilise the AI feedback? Share your experience.
3. What feedback points did you select, edit, deselect or add? Give examples and explain your rationale.

In the focus group interview, language instructors were also asked to share their views through a reflective marking journal designed to document their experiences with the AI-assisted marking platform. The template is organised into four sections. The first three sections focus on genre-based, referencing and language use, prompting instructors to provide overall impressions, identify areas for strengths and weaknesses, and suggest enhancements. The final section evaluates usability, effectiveness, and integration with the instructor's work, allowing for overall impressions and suggestions. At the end of the semester (week 13) after the language instructors finished marking the 662 students' writing, a total of 12 reflective journals from language instructors were collected and analysed. Finally, in Week 13, we collected 12 reflective journals to capture their thoughts on marking experiences while using the platform.

Table 1
Research Procedure and Data Collection

Date	Data	Number	Description
Week 5	Teacher feedback reports	14	Comparison between teachers' feedback reports against the original AI-generated feedback
Week 6	Focus group interview	1	A 1.5-hour interview about teachers' views on their experiences using the platform
Week 13	Reflective journals	12	Teachers' reflections on their marking experiences while using the platform

Through the triangulation of the data sources from the actual feedback outputs, user data from the AI feedback platform, interviews with the language instructors, and language instructor's reflective journals, this study examined language instructors' level of engagement with the AI-generated feedback and the impact of teacher agency on the collaborative feedback quality.

3.3 Coding schemes

Following a two-layer coding strategy, our data analysis provides a detailed examination of how AI action points are utilised. The two layers of code were documented in two Excel spreadsheets for descriptive analysis, enabling us to quantify how AI action points are utilised. The first layer includes tracking the percentages of points that are used, deleted, revised, or replaced with new suggestions through free text entries. The second layer focuses on specific actions taken by teachers in the "UR" (Revise with Revision) category. This coding reveals how teachers partially adopt AI feedback, highlighting the nuances of their decision-making and the adjustments made to enhance student writing. Together, these two layers of analysis offer valuable insights into the effectiveness and integration of AI feedback in language instruction.

Table 2
Coding of Teacher Utilisation of AI Feedback

Code	Description	Action taken by the language instructor
UE	Use exactly	Fully adopt the feedback point generated by AI verbatim
NU	Not Used	Decline the feedback point generated by AI
UR	Revise with revision	Partially adopt the feedback point generated by AI
FEC	Free Entry Comment	Add a new comment on an aspect not found in the AI-generated feedback points

Table 3
Coding of Teacher Actions within the UR Category

Code	Description	Action taken by the language instructor
QW	Quoting Students' Writing	Quote students' writing about the area that needs improvement
AE	Adding Explanation	Provide additional context or clarification about the feedback.
PE	Providing Example	Offer examples to illustrate feedback or suggestions.
ET	Explaining Writing Techniques	Describe specific writing techniques to improve student work.
DI	Deleting Irrelevant Information	Cut irrelevant/ inaccurate information
CD	Combining Different Points into One	Combine two or more feedback points generated by AI into one

3.4 Data analysis

Our analysis of teachers' feedback reports focuses on the distribution and comparison of different types of AI feedback utilisation. We analysed the percentages of AI action points categorised as used (UE), deleted (NU), revised (UR), or replaced (FEC). Additionally, we conducted a comparative analysis of teachers' engagement levels, examining the percentage of AI feedback utilised across teachers, categories and feedback types. This analysis allows us to triangulate the feedback data with insights from interviews and reflective journals, enabling us to further explore the factors facilitate or hinder teacher engagement with GenAI feedback.

The interview and reflective journal data were analysed using inductive analysis principles (Merriam & Tisdell, 2015). After transcribing the interviews, we reviewed the raw data and categorise it thematically, focusing on teachers' choices regarding AI feedback, factors that facilitate or hinder its use, and their perceptions. We started with loosely connected ideas for broad exploration and iteratively refine our coding by merging similar codes and addressing discrepancies. This non-linear process involved revisiting earlier steps as new insights arise, ensuring our analysis remains responsive to the data's complexities. Codes, such as quality of AI feedback points, teacher AI feedback literacy, critical digital literacy, and teacher agency, emerged from our analysis.

4 Findings

In this section, we will present our findings based on the two research questions. Section 4.1 to 4.5 will analyse teachers' feedback reports and interviews to address Research Question 1 (RQ1). Section 4.6 and 4.7 will present findings related to Research Question 2 (RQ2), drawing from teachers' reflective journals and supported by their feedback and comments from the interviews.

RQ1. To what extent do teachers engage with AI-generated feedback in their marking practices?

4.1 Engagement spectrum across markers

Table 4
Engagement Spectrum across Markers

Marker	UE	UR	FEC	Total number of AI feedback used	Rate of utilisation	Remark
11 markers	30	29	37	96	61.5%	Collaborative markers
Marker 3	5	0	0	5	100%	AI-Reliant Marker
Marker 12	6	0	0	6	100%	AI-Reliant Marker
Marker 6	0	0	4	4	0%	Sceptical Marker

As shown in Table 4, the 14 markers were categorised into three types based on how they used AI feedback. As shown in Table 4, the rate of AI utilisation varies significantly among the 14 markers. Three markers did not engage effectively in critical evaluation of AI feedback. Markers 3 and 12 utilised 100% of AI points without modifications, while Marker 6 completely rejected all AI feedback. Further analysis reveals that one of the AI points used by Marker 12 was considered inaccurate by other markers, suggesting the marker's lack of critical engagement and an over-reliance on AI-generated suggestions. In contrast, Marker 6 rejected all AI points entirely. During the interview, Marker 6 expressed a preference for traditional marking methods due to her unfamiliarity with the AI tool and her doubts about its effectiveness. The overall utilisation rate of the remaining 11 markers is 61.5%, indicating a balanced approach between AI and the markers. The results show that both capacity and willingness are essential for teachers to exercise agency in harnessing the affordances provided by the AI feedback platform. Teachers who view AI feedback positively because of their willingness to provide useful feedback and who possess critical digital literacy are more likely to act on these affordances. In contrast, teachers with lower willingness may completely reject the use of AI feedback, while those with high willingness but low capacity may over rely on AI feedback without critical evaluation of the writing sample. The markers' choices reflect what Toncelli and Kostka (2024) describe as a 'love-hate relationship' among faculty toward GenAI, with some instructors enthusiastically embracing AI tools for teaching, while others expressing strong resistance and scepticism.

4.2 Unbalanced utilisation of AI-generated feedback across categories

Table 5
Utilisation of AI-Generated Feedback in Category

AI Feedback Category	Commonly Utilised Feedback	Rarely Utilised Feedback
Genre-based Issues	1.2 Thesis statement (78.5%)	1.1 Word count (0%)
Referencing	2.4 Missing in-text citation (64.3%)	2.3 Lack of paraphrasing (0%)
Language use	3.3 Spelling mistakes (50%)	3.4 Fragment (14.3%)

Table 5 reveals that teachers use AI-generated feedback selectively, with some points receiving significantly more attention than others. For example, feedback on thesis statements is frequently utilised (78.5%), while feedback on word count is never used (0%). This selective usage indicates that teachers

perceive some feedback points as more valuable for student improvement than others. The commonly utilised points indicates that they may align with educators' primary concerns in the assessment process. However, the limited utility of certain feedback points raises questions about their relevance to the writing task. The findings show that the interaction between teacher agency and their engagement with AI feedback was shaped by the dynamic of teachers' AI readiness (i.e., capacity and willingness) and AI feedback output factors (e.g., features and quality of the feedback). In section 4.6 and 4.7, we will use interview data to further illustrate the factors that encourage or discourage the adoption of specific AI feedback points, providing deeper insights into this selective approach.

4.3 Types of teacher feedback assisted by AI tool

Table 6
Types of Feedback Utilisation

Feedback Type	Number of Action Points	Percentage of Utilisation
Use Exactly (UE)	41	36.9%
Not Used (NU)	126	0%
Use with Revision (UR)	29	26.1%
Free Entry Comment (FEC)	41	36.9%
Total Utilised Action Points	111	-
Average Action Points per Marker	7.9	-

According to Table 6, the total number of utilised action points was 111, with an average of 7.9 action points per marker. The UE type accounted for 41 action points, representing just 36.9% of the feedback. Additionally, 29 action points were classified as UR, making up 26.1% of utilised feedback after modification. FEC comprised 41 action points, accounting for 36.9% of total utilisation.

The significant number of unused action points (126 points) is understandable, given that teachers are expected to select only 4 to 8 points from the 14 GAI-recommended ones. The combined total of UR and FEC points constitutes 63% of the utilised feedback, indicating that most teachers did not simply copy the AI's recommendations verbatim. Instead, they exercised their judgment and expertise in marking, demonstrating critical engagement with the AI-generated feedback. By selectively incorporating relevant points, revising others, and even adding new feedback not recommended by the AI, teachers show both critical thinking and an awareness of the limitations of AI in the marking process.

Teachers engaged with the action points generated by AI feedback in distinctive ways, depending on their perception about the feedback quality. Some feedback points were utilised directly without revision, while others underwent revisions to enhance their effectiveness. Additionally, there were instances where teachers sought out new points based on the feedback provided. This variability indicates that the actual utilisation of the feedback platform's affordances relies on teachers' agency. Direct use of the feedback demonstrates a full realisation of its generative functionality. The practice of feedback revision denotes that teachers recognise the potential benefits of the feedback, but they exercise criticality to seek new opportunities for improvement. However, additional comments suggest that some teachers perceive no affordances or feel a need for enhanced affordances from AI feedback.

4.4 Revision of AI feedback in teacher marking

This section examines how teachers revise AI feedback under the Use with Revision (UR) category. The analysis of the 14 markers' feedback reports identifies six types of revision: quoting students' writing

(QW), adding explanation (AE), providing example (PE), explaining writing techniques (ET), deleting irrelevant information (DI), and combining different points into one (CD). It is important to note that some action points demonstrate more than one type of revision, such as combining QW and AE.

Table 7
Types of Feedback Revision

Types of revision	Total	Genre-based	Referencing	Language use
QW	10	5	1	4
AE	10	5	1	4
PE	4	4	0	0
ET	6	6	0	0
DI	4	0	3	1
CD	8	1	7	0
Total	42	21 (50%)	12 (28.6%)	9 (21.4%)

Although there are 29 UR feedback points, some feedback points contain multiple types of revisions. As a result, the total number of revisions identified across the 29 feedback points is 42, indicating that teachers perceive limitations in the AI comments, prompting them to elaborate further on the feedback provided. There is a notable focus on the genre-based category, which received 50% of the total revisions. The presence of the ET type exclusively within the genre-based category indicates that teachers recognise a close relationship between writing and specific genre types. In the referencing category, 7 out of 12 revisions were classified as CD, indicating that teachers found the AI's recommended comments repetitive. The language use category, which accounted for 9 revisions, includes instances where teachers quoted students' writing and clarified grammar rules. This highlights the teachers' efforts to connect the feedback to actual student work, enhancing its applicability.

In the following section, we present feedback from Marker 11 along with the comments from the interview.

Table 8
Revised Teacher Feedback

Teacher feedback	Teachers' comments in the interview
A thesis statement is a sentence that summarises the main point of your essay. It is not an entire paragraph, but it usually comes near the end of your introduction and tells the reader what you will argue and how you will justify your opinion. A good thesis statement should be concise, contentious and coherent.	AI Point 1.2 highlights that the student's essay lacks a thesis statement, which is a critical issue. It explains the importance of using a thesis statement to guide the argument. However, the student's writing indicates a lack of understanding of the basic concept of a thesis statement. Therefore, I added an explanation to assist him.

While the AI feedback highlights a critical issue, Marker 11 observes that the student does not fully understand what a thesis statement is, prompting her to offer an explanation to assist the student. This reflects that the teacher acknowledges the benefits of the feedback and looks for new affordance for improvement based on her awareness of the student's level of understanding and her commitment to providing helpful feedback. Further comparison between AI and teacher feedback indicates that AI

primarily focuses on evaluating the essay itself. In contrast, human teachers take into account additional factors, such as the learner's level of understanding, the specific genre-related requirements of the assignment, and the development of writing skills through broader planning strategies.

4.5 Free entry action points

Table 9
Free Entry Action Points

Free entry feedback points		%
1	Opening and closing sentences of feedback report	64.3%
2	Unclear introduction	71.4%
3	Topic sentence	50%
4	Background of the book	21.4%
5	Inconsistent verb tense	12.2%
6	Lack critical analysis	21%
7	Cohesion	7.8%
8	Tone and style	14.1%
9	Mechanical headings	14.1%
10	Link ideas	18%
11	Others: Title Coherence Double brackets punctuation Avoid translation software Read the checklist and assignment guidelines carefully Provie resources or tools about APA	16.4%

Table 9 presents 16 free-entry feedback points from markers, highlighting various areas of focus in their evaluations. This feedback identifies weaknesses in specific parts of the essay, such as the introduction and organisation. Some comments extend beyond essay evaluation, offering advice on completing writing tasks, such as reminding students to carefully read the assignment requirements. Additionally, there are ethical considerations, such as avoiding translation software. AI feedback primarily focuses on technical aspects of writing. In contrast, teacher feedback takes a broader perspective, incorporating personalised advice, ethical considerations, and strategies for improvement, promoting a more holistic approach to writing development.

The highest percentages of comments addressed unclear introductions (71.4%) and the opening and closing sentences of the feedback report (64.3%). The consensus among the markers indicates that most teachers perceived that these two specific areas were inadequately addressed by the AI platform.

In the following, we present one feedback point and analyse teacher's rationale for the use of additional action points.

> Inconsistent verb tense: When you summarise the book, you could use present tense or past tense, and it would be more desirable to ensure consistency in tense usage. I notice that the tense changes unnecessarily in parts of your book summary. [Marker 2, feedback report]

Compared to AI feedback points in language use category, which only address the surface errors, Marker 2 identified a systematic error that AI overlooked. This observation is more valuable because it highlights a recurring error rather than a minor spelling mistake. Teachers noted in interviews that tense errors are common among students, suggesting they consider the frequency of errors and their impact on essay effectiveness. In contrast, AI tends to randomly select errors without addressing broader patterns.

RQ2: How does teacher agency influence human-AI collaborative feedback decisions?

In the following sections, we first analyse data from teachers' reflective journals to understand their reasons for accepting or rejecting AI-generated feedback. Next, we will triangulate this self-reported data by examining how teachers incorporate AI feedback into their reports and their preferences for specific feedback points. We will also draw on interview data to further clarify their feedback practices.

4.6 Encouraging factors for using AI feedback

Table 10

Selective Comments on Using AI Feedback (Reflective Journal)

Reasons for adapting the AI feedback
The genre-based feedback is very comprehensive.
I think the AI interface is very useful and effective. This platform is user-friendly, with clear guidelines on its usage. It saves my time.
The AI platform helps me with word choice and expression when I have occasional struggles, freeing up time for me to focus on professionally evaluating student writing.
The platform is user-friendly and integrates seamlessly with my teaching workflow. Its intuitive design makes it easy to navigate, and the effectiveness of the AI-generated feedback has been impressive.
I like how it would capture several instances from the student's script as part of the explanation.
Sometimes, the AI-generated feedback can help me notice the "blind spot" that I may not realize while reading the students' first draft.
I love this idea of giving the instructors the option to select / deselect specific, individual feedback points, and to add in our own.

Teachers' reflection shows that they adapted AI feedback for several compelling reasons, demonstrating high levels of engagement. First, the AI feedback helped reduce the teachers' cognitive demands associated with feedback decisions by: (1) providing comprehensive genre-specific feedback, (2) saving time on word choice and expression, allowing teachers to focus on evaluating higher-order writing issues, (3) highlighting specific instances in student writing to enhance clarity, and (4) identifying blind spots that teachers previously overlooked and improving the overall feedback. In addition, teachers appreciated the user-experience design, noting that it saves time and integrates seamlessly into their workflow. The flexibility to select or delete individual feedback points and add personal comments further enhances the

adaptability of the AI feedback, allowing teachers to tailor their responses to better meet their students' needs.

In summary, the combination of AI's genre-specific feedback, user-friendly AI interface design, and adaptable feedback mechanisms made AI a supportive partner and contributed to teachers' active engagement with AI tools in their professional assessment practice.

As shown in Table 5 in Section 4.2, AI action point 2.4 about missing in-text citation is well accepted by markers, with a rate of 64.3%. To further explore why this action is favoured, we have examined an actual feedback report from one of the markers. Below are Marker 8's feedback report and interview comments, as Marker 8 applied this point exactly.

> 2.4 Missing in-text citations: The essay lacks in-text citations. For example, the sentence "The film reflects the phenomenon of bullying in adolescence and reflects the phenomenon that some children lack attention due to the lack of energy of their parents, which is common in families with many children." needs an in-text citation to support the argument. Please refer to the APA 7th edition referencing style guide for more information on how to include in-text citations: https://www.polyu.edu.hk/elc/independent-learning/language-resources/apa/ [Marker 8, feedback report].

Marker 8 expressed her comment on AI point 2.4.

> Item 2.4 is a valuable AI-generated feedback point that aligns with my views on writing requirements by stressing the importance of proper citation practices. It identifies the issue of missing in-text citations with a specific example, guiding my students on where to improve. Additionally, it provides a link to the APA referencing style guide, ensuring they have the resources needed to meet academic standards and maintain integrity in their writing. [Marker 8, interview]

The analysis indicates that the AI-generated feedback point on missing in-text citations is a significant reason why the marker chooses to utilise it. This feedback aligns closely with the teacher's understanding of task requirements about the importance of proper citation practices. Additionally, the AI point provides concrete examples for improvement, which guides students in identifying specific areas that require attention. Furthermore, the inclusion of resources such as web links and study tools enhances the feedback's effectiveness by equipping students with the necessary information to improve their writing. The marker values these qualities highly, believing that this feedback will facilitate meaningful enhancements in the student's academic writing skills. In this instance, the distributed agency between the teacher and AI, due to the interaction between the teacher's capacity and willingness on the one hand and features of AI feedback on the other, is clearly evident: the AI identified an issue that the teacher also deemed important. The teacher then applied their critical judgment and feedback literacy to evaluate the AI-generated feedback.

4.7 Discouraging factors for rejecting AI feedback

Table 11 reveals teachers' concerns about the effectiveness of AI feedback in evaluating student writing. Firstly, teachers believe the AI system treats all areas with equal priority and struggles to identify core issues in students' work. They argue that minor errors, such as spelling, should be deprioritised in favour of more significant feedback on overall organisation and structure. This even distribution of focus leads to repeated points in the referencing category. This result aligns with Guo and Wang's (2023) finding that ChatGPT feedback tends to be evenly distributed across categories, whereas teacher feedback often prioritises specific aspects of student writing. Teachers' comments indicate that they were able to identify limitations in the AI feedback mechanism and applied their critical digital literacy (Darvin, 2025) to evaluate the AI feedback effectively.

Additionally, AI is perceived as lacking contextual understanding of the assignments. For instance, it fails to recognise that students are expected to focus on building their main argument in the current draft and to reinforce it with secondary sources in the next draft. Sometimes, AI feedback is overly technical, suggesting a misunderstanding of students' English level, particularly since most students have an English proficiency around IELTS 5.5. This finding aligns with Fan et al.'s (2024) study, which indicate that students perceived teacher feedback as more effective than AI-generated feedback for genre-specific assignments.

Overall, teachers seek a more tailored approach from AI that identifies specific areas for improvement without relying solely on predetermined categories, underscoring the need for a more sophisticated and contextual understanding of student writing. This reflects the design dimension of teacher feedback literacy (Carless & Winstone, 2020), wherein feedback literature teachers actively enacted their agency by adapting and refining feedback to better align with pedagogical goals and student needs.

Table 11

Selective Comments on Using AI Feedback (Reflective Journal)

Reasons for rejecting AI feedback
It appears that AI-assisted feedback platform cannot always detect the core issues of students' writings.
Many incorrect or repeated points are found in the reference category.
The system should provide more suggestions for the overall organisation or structure.
The human teacher understands the writing context better.
The feedback about using secondary sources should appear in the second draft, not the first draft.
Occasionally, the language feedback can be overly technical, which might be challenging for some students to understand.
I would like to know if the system can identify specific areas where an essay needs improvement without relying on the predetermined areas.

As shown in Table 5 of Section 4.2, all markers rejected AI action points 1.1 on word count and 2.3 on insufficient paraphrasing, leading to a utilisation rate of 0%. To better understand the reasons for this rejection, we compared the data in Table 11, which includes AI feedback points and teacher comments, to further explore the issue.

The AI emphasises strict adherence to word count for conciseness, while markers priortise content quality over numerical limits, highlighting differing views on the writing process. Marker 4 rejected the AI feedback regarding word count and instead focused on the lengthy introduction, emphasizing the need to improve both its structure and length. The teacher recommends a more structured approach to enhance coherence and clarify the argument, helping the student achieve better organisation in the essay. This perspective values depth and quality over rigid number of words, illustrating a shift from the AI's emphasis on format to a focus on content in writing assessment. It highlights the need for comprehensive planning and critical engagement, exposing a gap in the AI's feedback that overlooks the importance of depth in essay writing. Teachers' feedback literacy and willingness to exercise agency to perceive and

evaluate the affordance provided by AI feedback enabled them to prioritise certain writing issues over others, with the aim of designing effective feedback that genuinely supports student learning.

In point 2.3, the AI emphasises conventional academic writing by advocating for limited direct quotation and strong paraphrasing. However, it fails to consider the specific genre of a book review, where quoting is essential for analysis. Marker 11 acknowledges the importance of quotes in this context, suggesting a flexible approach that values evidence, in contrast to the AI's strict adherence to general academic standards. This example demonstrates that AI does not fully grasp the specific requirements of the genre, highlighting the indispensable role of the teacher in genre-based feedback practices.

Table 12

Reasons for Rejecting AI Feedback (Insight from Interview)

AI feedback point	Teacher's comment in the interview
1.1 The essay exceeds the required word count of 700 words. It is important to adhere to the word limit as it demonstrates your ability to express your ideas concisely. Exceeding the word limit may lead to unnecessary details and may lose the reader's interest. Please revise your essay to meet the word limit.	A lengthy introduction limits space for body paragraphs, affecting the paper's effectiveness. A more structured approach would enhance coherence in his writing and clarify his argument, guiding him toward better organisation in his essay. [Marker 4, interview]
2.3 The essay lacks proper paraphrasing and contains direct quotes without citation, such as, "When given the choice between being right or being kind, choose kind." Effective paraphrasing demonstrates a deeper understanding of the text.	The lack of paraphrasing is noted, but the book review should include evidence from the text, allowing for some direct quotes. Therefore, I would remove point 2.3 about paraphrasing from my marking. [Marker 11, interview]

5 Discussion

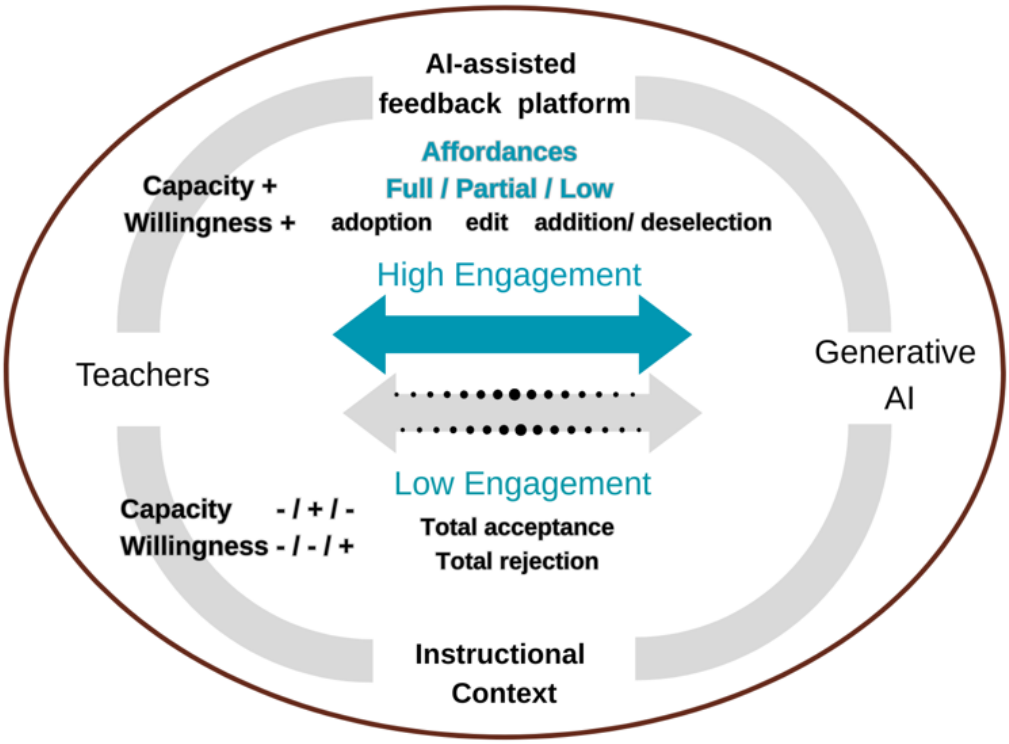
5.1 A model of teachers' behavioural engagement with GenAI feedback

This study has sought to explore teachers' behavioural engagement with AI-generated feedback in their marking practices and the influence of teacher agency on their engagement. Based on the findings, a model of teachers' behavioural engagement with GenAI feedback has been developed (Figure 2). The model depicts teachers' different modes of behavioural engagement with GenAI feedback, depending on whether they exercise agency to perceive affordances in such feedback, and if so, the various degrees of perceiving and shaping affordances as determined by teacher capacity and willingness. This model highlights the relational feature of agency (Biesta et al., 2015), regarding the influence of teacher agency on teacher engagement with AI feedback as the interaction between teacher factors (i.e., capacity-related factors and willingness-related factors) and AI feedback-related factors (e.g., features of AI feedback). In the instructional context, teachers may demonstrate a low or high level of engagement with AI feedback. When teachers do not exercise agency to perceive affordances in AI feedback because they are incapable of collaborating with AI to provide helpful feedback to students and/or unwilling to do so (as represented by "Capacity- & Willingness-", "Capacity+ & Willingness-" and "Capacity- & Willingness+"), they may display a low level of engagement (as represented by the double-headed dotted line connecting Teachers and Generative AI) by totally accepting or rejecting AI feedback (as represented by "Total acceptance" and "Total rejection"). In contrast, when teachers exercise agency to perceive affordances in AI feedback based on their capacity and willingness (as represented by "Capacity +" and "Willingness +"), they may perceive various degrees of affordances in AI feedback (as represented by "Full", "Partial" and "Low")

and the need to shape new affordances for students in the cases of partial and low affordances. They will then display a high level of engagement (as represented by the double-headed solid line connecting Teacher and AI feedback) by using exact AI feedback (as represented by "adoption" in the model), using such feedback with revision (see "edit") or offering additional comments (see "addition/deselection"), the latter two of which represent the shaping of new affordances for students. The degree to which affordances are perceived hinges on the extent of alignment between teacher factors (i.e., capacity-related and willingness-related factors) and Generative AI (e.g., features of GenAI feedback).

Figure 2

Model of Teachers' Behavioural Engagement with GenAI Feedback



5.2 Teacher engagement with AI-generated feedback

RQ1 explored the extent to which teachers engaged behaviourally with AI-generated feedback in their marking practices. We have identified three types of markers: 1 AI-sceptical, 2 AI-reliant and 11 cooperative teachers. The AI-sceptical and AI-reliant markers showed a low level of engagement with AI feedback. They either totally rejected AI suggestions and provided entirely new feedback or adopted the AI feedback without any modifications despite its inaccuracy. In contrast, the 11 markers displayed a high level of engagement with AI feedback. They not only critically evaluated AI feedback and used it selectively, but also enhanced AI feedback with revisions (e.g., providing personalised comments and explaining writing techniques) or even provided additional comments. Most revisions of AI feedback and additional comments fell within the genre-specific category, aligning with Fan et al.'s (2024) findings, which highlight the importance of genre-specific feedback from teachers. The various forms of low and high levels of teacher engagement with AI-feedback suggest that teachers' responses to GenAI feedback were diverse, revealing a spectrum of behavioural engagement. Previous research has mainly focused on comparing features of teacher feedback and GenAI feedback (Fan et al., 2024; Guo & Wang, 2023; Steiss et al., 2024). This study adds to previous research by investigating the under-researched topic of teacher engagement, revealing various forms of teacher behavioural engagement with GenAI feedback in teacher-AI collaborative feedback practices.

5.3 The influence of teacher agency on teacher engagement

RQ2 investigated the influence of teacher agency on teacher behavioural engagement with GenAI feedback. From an ecological perspective (Gibson, 1986; van Lier, 2000), such an influence is conceptualised as the interaction between teacher factors (i.e., capacity-related factors and willingness-related factors) and GenAI -related factors (e.g., features of AI feedback). This study shows that the low or high level of teacher behavioural engagement may be affected by the degree to which teachers exercised agency to collaborate with GenAI based on their capacity and willingness. For example, the AI sceptical marker rejected all AI feedback and provided her own feedback to the students because she was unfamiliar with the AI tool (i.e., a lack of capability to use the AI tool as part of critical digital literacy) and doubted the effectiveness of AI feedback (i.e., biased beliefs about AI feedback). Due to a lack of capability and willingness (see "capacity-" and "willingness-" in Figure 2), she did not exercise agency to interact with AI feedback to perceive its affordances, showing a low level of engagement through total rejection of AI feedback (see "Low Engagement" and "Total rejection" in Figure 2). In the case of AI-reliant markers, they also did not exercise agency to fully interact with AI feedback to perceive its true affordances because of the inability to detect problematic AI feedback despite their willingness to provide helpful feedback (see "Capacity-" and "Willingness+" in Figure 2). As a result, they displayed a low level of engagement by total acceptance of AI feedback without careful consideration (see "Low Engagement" and "Total acceptance" in Figure 2). It is also likely that a lack of willingness on the part of the markers may prevent them from engaging actively with AI feedback despite their capacity (see "Capacity+" and "Willingness-"), although this was not shown in the findings.

In the study, the 11 markers demonstrated a high level of engagement because they were both able and willing (see "Capacity+" and "Willingness+" in Figure 2) to perceive various degrees of affordances in AI feedback as a basis for choosing the different modes of active engagement. The various extents to which teachers were able to perceive affordances depend on the alignment between teacher factors (i.e., capability- and willingness-related factors) and GenAI-related factors (e.g., features of AI feedback). For the engagement mode of "Adoption", Marker 8's favourable comment on AI point 2.4 is a typical example. The features of the AI feedback (i.e., addressing task requirements and reflecting features of effective feedback, which includes extra resources) aligned with the teacher's knowledge of assignment requirements, teacher feedback literacy (Lee, 2021) (e.g., teacher understanding of high-quality feedback) as well as his or her willingness to collaborate with AI to provide helpful feedback to students. The full alignment between teacher factors and AI feedback-related factors enabled Marker 8 to perceive the full affordance of AI point 2.4 and then adopt it fully for feedback provision (See "Full affordances" and "Adoption" in Figure 2). The facilitating factors for using AI feedback identified in the study also imply a match between teacher factors and GenAI-related factors, suggesting teachers' active role in perceiving the full potential of AI feedback for adoption. The teacher factors include teacher feedback literacy regarding understanding of effective teacher feedback such as the importance of comprehensive feedback in general and genre-based feedback in particular, critical digital literacy concerning teacher ability to critically use digital tools (Darvin, 2025), and teacher motivation to work with AI to produce helpful feedback, while GenAI-related factors include features of AI feedback and user-friendly AI interface.

For the engagement mode of "Edit", the teachers perceived partial affordances in AI feedback and the need to shape new affordances (see "Partial affordances" and "Edit" in Figure 2) due to the partial alignment between teacher-related factors and GenAI-related factors, leading to the teacher action of using AI feedback with revision. One typical example is Marker 11's adaptation of the AI feedback on thesis statement. On the one hand, Marker 11 perceived the affordance in the AI feedback on thesis statement based on her knowledge of the genre-specific assignment requirements and her motivation to provide helpful feedback. On the other hand, however, there was misalignment between the teacher's knowledge of students' level of understanding and what the AI feedback can afford (i.e., a lack of

explanation about thesis statement to facilitate student understanding). Given the partial alignment, the teacher shaped new affordances by adding an explanation about thesis statement according to her professional knowledge and the commitment to providing helpful feedback.

Concerning the engagement mode of "Addition/Deselection", the teachers perceived low affordances in AI feedback, and they realised the need to shape new affordances in their own feedback (see "Low affordances" and "Addition/Deselection" in Figure 2) because of the misalignment between teacher factors and GenAI-related factors. First, the misalignment caused the teachers to deselect AI feedback. For example, Marker 4's emphasis on content issues over word limits (i.e., teacher feedback literacy) and willingness to collaborate with AI to offer students helpful feedback led her to critically evaluate AI feedback and see low affordances in the AI feedback that highlighted the importance of word limit for conciseness. Similarly, Marker 11's knowledge of genre requirements (i.e., the essence of quoting for analysis in a book review) and willingness to provide helpful feedback misaligned with the general AI feedback on a lack of paraphrase, which did not take into consideration the specific genre of a book review. Such a misalignment led to the teacher's perception of low affordances in AI feedback, thus deselecting AI feedback. The discouraging factors for rejecting AI feedback also indicate a misalignment between teacher factors (e.g., teacher feedback literacy regarding teacher understanding of effective feedback, teacher understanding of genre-specific assignment requirements and teacher understanding of student factors such as students' proficiency level) and GenAI-related factors (e.g., features of AI feedback). In other words, the aforementioned teacher factors prevented teachers from seeing the learning potential of AI feedback, given its characteristics. Second, the perception of low affordances made the teachers realise the necessity of shaping new affordances in additional teacher comments. A case in point is Marker 2's practice. His understanding of students' common errors (i.e., inconsistent tense use) and willingness to provide useful feedback misaligned with the random surface errors identified by AI, causing him not only to deselect such AI feedback, but also to add his own feedback on the common error of inconsistent verb tense to facilitate students' writing improvement. It is important to note that even in the cases of low affordances perceived in AI feedback, it does not mean that teachers did not exercise agency--they still exercised agency to shape new affordances (Kyttä, 2004) by adding their own comments for students to use.

In short, the influence of teachers' agency on their behavioural engagement is manifested through the different extents to which their capacity and willingness enabled them to take the initiative to perceive learning opportunities of AI feedback as a basis for collaborating (or not collaborating) with GenAI for feedback provision. Previous research (Fan et al., 2024; Guo & Wang, 2023; Steiss et al., 2024) has not explored teacher engagement with AI feedback as well as factors affecting such engagement. This study represents one of the few attempts to explore these two issues and highlights the influence of teacher agency on teacher engagement with GenAI feedback. In particular, this study emphasises the relational aspect of agency (Biesta et al., 2015) in the context of teacher-AI collaboration and proposes a model underscoring the interactional impact of teacher factors and GenAI-related factors on teacher engagement with AI feedback. This relational aspect of agency (Biesta et al., 2015) reflects the distributed agency (Jones, 2022) involving both teachers and AI tools.

6 Implications

5.1 Implication for teacher-AI collaboration in feedback practices

This study provides pedagogical implications for teachers utilising AI feedback, emphasising the need for a critical and balanced approach. AI and teachers possess complementary strengths in feedback practices. AI tools can significantly lighten teachers' workloads, offer comprehensive feedback, and provide tailored resources to enhance student learning (cf. Barrot, 2023a, 2023b; Kohnke et al., 2023).

However, there are limitations: some AI-generated feedback may be inaccurate, overlook key aspects (cf. Guo & Wang, 2023), and fail to address individual student needs (cf. Barrot, 2023a, 2023b).

As AI has been significantly transformed English Language Teaching (ELT), training focused on critical engagement with AI feedback becomes essential for professional growth. Strengthening AI feedback literacy, in addition to general feedback literacy and digital literacy (Mullen, 2025), can lead to more effective marking practices. Without adequate training to support their understanding and reflection, sceptical educators, like Marker 6, may continue to view AI as peripheral to their professional identity. Conversely, AI-reliant teachers, such as Marker 3 and Marker 12, may overlook opportunities to critically evaluate the feedback they receive, which could undermine its effectiveness.

With proper training, teachers can effectively collaborate with AI to provide high-quality writing feedback (Barrot, 2023a, 2023b). They can bring essential contextual understanding, genre-specific insights, and personal connections that foster trust and motivation among students. Their expertise in identifying core issues and delivering holistic evaluations makes their role indispensable, especially in discipline-specific contexts.

Critical evaluation would empower educators to determine when to rely on AI and when to intervene, thereby safeguarding the quality and effectiveness of their feedback. Striking a balance between leveraging the benefits of AI and addressing its challenges is crucial for optimising feedback practices.

5.2 Implication for AI training

This study offers implications for schools, educational institutions, and organisations developing customised AI feedback tools. Our findings, along with previous research, highlight several weaknesses in AI feedback systems. For instance, ChatGPT often distributes feedback evenly across various aspects of essays (Guo & Wang, 2023), which can sometimes obscure critical issues that require attention.

To enhance the effectiveness of AI-assisted feedback, when training AI tools, it is essential to shift the focus from evaluating local-level issues, such as spelling and grammar errors, to a more holistic evaluation of essays, examining overall structure and the construction of arguments. Training the AI with contextual information about students' levels and expectations will enable it to deliver more personalised feedback, fostering a stronger emotional connection with learners. Incorporating exemplars from diverse proficiency levels will help the AI better understand and respond to the unique needs of students with different performance.

Moreover, focusing on genre-specific materials during training will improve the AI's sensitivity to genre-related elements, ensuring that feedback is both appropriate to the genre, relevant to the disciplines and constructive in supporting students' active learning and facilitating their approximations as members of their disciplinary communities. Additionally, addressing systematic errors by training the AI to prioritise feedback based on the frequency and significance of these errors will result in more meaningful assessment practices. By implementing these strategies, AI can become a more effective tool for scaffolding writing practices, supporting student writing development, ultimately transforming the feedback process in AI-supported educational settings.

7 Conclusion and Limitation

This study analyses the interactions of 14 teachers with an AI-assisted feedback platform specifically designed to evaluate discipline-specific essays that adhere to genre conventions. The results reinforce existing research on the importance of teacher feedback literacy and distributed agency between teachers and the AI tool. Teachers' agency is influenced by both willingness-related and capacity-related factors, such as teacher beliefs, teacher feedback literacy, and critical digital literacy. These factors collectively

shape teachers' engagement with AI feedback. A significant contribution of this study is the proposal of a conceptual model, which represents varying levels of engagement resulting from the interaction between teachers and AI agency.

Despite its contributions, this study has several limitations. This study analysed how different teachers mark one single paper. Using multiple markers provides advantages, such as diverse perspectives that enhance our understanding of how AI feedback is perceived and applied, revealing a spectrum of engagement. Future studies could further benefit from analysing the marking of multiple essays. This approach would enable researchers to examine the types of feedback provided and determine whether certain feedback is more prevalent for different student groups, such as high achievers versus low achievers. Analysing multiple essays can also identify common mistakes, highlighting recurring challenges in writing.

Additionally, future research could compare the human-AI collaborative assessment of different genre types to deepen the understanding of how genre influences teachers' marking engagement. Analysing specific genre types across samples could yield insights into disciplinary genre-specific concerns. Lastly, the relatively small sample size may limit the generalisability of the findings, so future research should involve larger samples.

Funding Statement

This project, running from 2022 to 2025, is funded by the Quality Enhancement Support Scheme (QESS) under the Self-financing Post-secondary Education Fund of the Education Bureau of the HKSAR Government. The total funding received is HK$2,299,996, and the project number is 03/QESS/2022.

References

Ahearn, L. M. (2001). Language and agency. *Annual Review of Anthropology, 30*(1), 109-137. https://doi.org/10.1146/annurev.anthro.30.1.109

Barrot, J. S. (2023a). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning, 36*(4), 584-607. https://doi.org/10.1080/09588221.2021.1936071

Barrot, J. S. (2023b). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, Article 100745. https://doi.org/10.1016/j. asw.2023.100745

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, *21*(6), 624–640. https://doi.org/10.1080/13540602.2015.1044325

Biesta, G., & Tedder, M. (2007). Agency and learning in the life course: Towards an ecological perspective. *Studies in the Education of Adults*, *39*(2), 132–149. https:// doi.org/10.1080/02660830.2007.11661545

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education, 28*(1), 150-163. https://doi.org/10.1080/13562517.2020.1782372

Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, *83*, 101395.

Cheng, X., Liu, Y., & Wang, C. (2023). Understanding student engagement with teacher and peer feedback in L2 writing. *System*, 119, 103176.

Cong-Lem, N. (2024). Teacher agency for change and development in higher education: A scoping literature review. *International Journal of Educational Reform*, *0*(0). https://doi.org/10.1177/10567879231224744

Cunningham, K. J., & Link, S. (2021). Video and text feedback on ESL writing: Understanding attitude and negotiating relationships. *Journal of Second Language Writing*, *52*. https://doi.org/10.1016/j.jslw.2021.100797

Darvin, R. (2025). The need for critical digital literacies in generative AI-mediated L2 writing. *Journal of Second Language Writing, 67*, Article 101186. https://doi.org/10.1016/j.jslw.2025.101186

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*(2), 335–349.

Fan, Y., Tan, S., & Lim, G. Y. W. (2024). EAP teacher feedback in the age of AI: Supporting first-year students in EFL disciplinary writing. *Australian Journal of Applied Linguistics, 7*(3), 1943. https://doi.org/10.29140/ajal.v7n3.1943

Gibson, J. J. (1986). *The ecological approach to visual perception*. Lawrence Erlbaum.Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *Language, Learning and Technology, 28*(2), 5-30.

Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology, 28*(2), 5-30. https://doi.org/10.64152/10125/73570

Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies, 29*(7), 8435-8463.

Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System, 126*, 103502. https://doi.org/10.1016/j.system.2024.103502

Han, T., & Sari, E. (2024). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning, 37*(4), 961-985. https://doi.org/10.1080/09588221.2022.2067179

Han, Y., & Reinhardt, J. (2022). Autonomy in the digital wilds: Agency, competence, and self-efficacy in the development of L2 digital identities. *TESOL Quarterly, 56* (3), 985-1015. https://doi.org/10.1002/tesq.3142

Huang, J., & Teng, M. F. (2025). Peer feedback and ChatGPT-generated feedback on Japanese EFL students' engagement in a foreign language writing context. *Digital Applied Linguistics, 2*, 102469-102469.

Hyland, F. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research, 4*(1), 33-54.

Hyland, K. (2004). *Disciplinary Discourses: Social Interactions in Academic Writing*. Longman.

Hyland, K. (2022). English for specific purposes: What is it and where is it taking us?. *ESP Today-Journal of English for Specific Purposes at Tertiary Level, 10*(2), 202-220.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, *39*(2), 83-101.

Jiang, L., Yu, S., & Wang, C. (2020). Second language writing instructors' feedback practice in response to automated writing evaluation: A sociocultural perspective. *System, 93*, Article 102302. https://doi.org/10.1016/j.system.2020.102302

Jones, R. H. (2022). Commentary: Critical digital literacies as action, affinity, and affect. *TESOL Quarterly, 56*(3), 1074-1080.

Kohnke, L., Moorhouse, B., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/ 0033688223116286

Kyttä, M. (2004). The extent of children's independent mobility and the number of actualized affordances as criteria for child-friendly environments. *Journal of Environmental Psychology*, *24*(2), 179–198. https://doi.org/10.1016/s0272-4944(03)00073-2

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.

Lee, I. (2019). L2 writing teacher education and feedback literacy development: Perspectives from Hong Kong. In L. Seloni, & S. Henderson Lee (Eds.), *Second language writing instruction in global contexts: English language teacher preparation and development* (pp. 31-51). Multilingual Matters.

Lee, I. (2021). The development of feedback literacy for writing teachers. *TESOL Quarterly, 55*(3), 1048-1059. https://doi.org/10.1002/tesq.3012

Li, Z. (2021). Teachers in automated writing evaluation (AWE) system-supported ESL writing classes: Perception, implementation, and influence. *System, 99*. https://doi.org/10.1016/j.system.2021.102505

Lyster, R., & Ballinger, S. (2011). Content-based language teaching: Convergent concerns across divergent contexts. *Language Teaching Research, 15*(3), 279-288. https://doi.org/10.1177/1362168811401150

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation.* John Wiley & Sons.

Moorhouse, B. L., Wan, Y., Wu, C., Wu, M., & Ho, T. Y. (2025). Generative AI tools and empowerment in L2 academic writing. *System*, *133*, 103779. https://doi.org/10.1016/j.system.2025.103779

Mullen, M. (2025). Equipping Language Teachers with the Digital Literacy to Support Student Learning in and Beyond Classrooms. *International Journal of TESOL Studies*, *7*(3), 106-129. https://doi.org/10.58304/ijts.250715

Nesi, H. & Gardner, S. (2006). Variations in disciplinary culture: University tutors' views on assessed writing tasks. In R. Kiely, G. Clibbon, P. Rea-Dickins, & H. Woodfield (Eds.) *Language, Culture and Identity in Applied Linguistics* (pp. 99-117). Equinox.

Park, J., & Ahn, S. (2022). L2 learners' cognitive and behavioral engagement with written corrective feedback. *English Teaching*, 77(3), 133-152. https://doi.org/10.15858/engtea.77.3.202209.133

Pearson, W. S. (2024). Affective, behavioural, and cognitive engagement with written feedback on second language writing: A systematic methodological review. *Frontiers in Education, 9,* 1285954. https://doi.org/10.3389/feduc.2024.1285954

Peungcharoenkun, T., & Waluyo, B. (2024). Students' affective engagements in peer feedback across offline and online English learning environments in Thai higher education. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(1). https://doi.org/10.1186/s40862-024-00286-w

Reed, E. S. (1993). The intention to use a specific affordance: A conceptual framework for psychology. In R. H.Wozniak, & K.W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 45-76). Hillsdale, NJ: Lawrence Erlbaum.

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction, 91*, Article 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Teng, M. F. (2019). *Autonomy, agency, and identity in teaching and learning English as a Foreign Language*. Springer.

Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education*. *Artificial Intelligence, 7*, 100270.https://doi.org/10.1016/j.caeai.2024.100270

Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, *60*(1), e12811. https://doi.org/10.1111/ejed.12811

Teng, M. F., & Huang, J. (2025). Incorporating ChatGPT for EFL writing and its effects on writing engagement: *International Journal of Computer-Assisted Language Learning and Teaching*, *15*(1), 1-21. https://doi.org/10.4018/IJCALLT.367874

Teng, M. F., & Yip, W. C. (2025). A narrative inquiry into language teacher identity construction in the global GenAI context. *International Journal of TESOL Studies, 250901*, 1-21. https://doi.org/10.58304/ijts.250901

Toncelli, R., & Kostka, I. (2024). A love-hate relationship: Exploring faculty attitudes towards GenAI and its integration into teaching. *International Journal of TESOL Studies, 6*(3), 77-94. https://doi.org/10.58304/ijts.20240306

Tong, E.K., Pun, C., & Siu, P. (2023). Enhancing disciplinary learning experience through an adjunct English-across-the-curriculum model. In T. Slater (Ed.). *Social practices in higher education: A knowledge framework approach to linguistic research and teaching* (pp. 382-409). Equinox.

Tsao, J., & Nogues, C. (2024). Beyond the author: Artificial intelligence, creative writing and intellectual emancipation. *Poetics, 102*, Article 101865. https://doi.org/10.1016/j.poetic.2024.101865

van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J.P. Lantolf (Ed.), *Sociocultural theory and second language learning: Recent advances* (pp.245-259). Oxford University Press.

Wingate, U. (2018). Academic literacy across the curriculum: Towards a collaborative instructional approach. *Language Teaching, 51*(3), 349-364. https://doi.org/10.1017/S0261444816000264

Zhang, B. (2021). Engaging in dialogue during collaborative writing: The role of affective, cognitive, and social engagement. *Language Teaching Research*, *29*(1), 33-62. https://doi.org/10.1177/13621688211054047

Zhang, Y., & Dikilitaş, K. (2025). Investing in critical digital literacies involving generative AI: A multiple-case study of second language novice teachers. *International Journal of TESOL Studies, 250626,* 1-25. https://doi.org/10.58304/ijts.250626

Zhang, Z. V., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. *Assessing Writing*, 58, 100784.

Zou, S., Guo, K., Wang, J., & Liu, Y. (2025). Investigating students' uptake of teacher-and ChatGPT-generated feedback in EFL writing: a comparison study. *Computer Assisted Language Learning,* 1-30. https://doi.org/10.1080/09588221.2024.2447279

Dr. Deng Chunrao Jessica is a Senior Lecturer in the Division of Languages and Communication at the Polytechnic University of Hong Kong. She obtained her Doctoral Degree from the Faculty of Education, The University of Hong Kong. Her research focuses on second language education, communicative teaching approaches, and formative feedback. Currently, she is the primary investigator on a government-funded project for an AI-assisted academic writing platform.

Dr Lin Shuyang Sylvia is a Senior Lecturer at the College of Professional and Continuing Education, The Hong Kong Polytechnic University. She obtained her Doctoral Degree from the English Department, City University of Hong Kong. Her research interests include academic writing, L2 motivational practices, critical thinking, and feedback for learning. She has published journal articles and presented at conferences in the fields of L2 motivational studies, critical thinking, and L2 feedback.

Dr. Tong Ka-man Esther is the Head of the Division of Languages and Communication and Director of the CPCE Centre for Pedagogic Research at the College of Professional and Continuing Education, The Hong Kong Polytechnic University. She serves as Co-Principal Investigator for a government-funded project on AI-assisted academic English writing and leads a SCOLAR-funded study exploring adjunct English instruction. Her research centres on AI-mediated language learning and CLIL assessment. Dr. Tong has been recognised for her contributions to educational practice, receiving awards such as the PolyU CPCE Dean's Award for Outstanding Teaching Performance/Achievement (2014/15) and the FSTE General Education Outstanding Teaching Award (2017).

Dr. Ma Jingjing Maggie is Associate Professor in the Department of English at the Hang Seng University of Hong Kong. Her research interests include formative assessment, teacher education, student feedback literacy and L2 writing. Her recent publications appeared in journals such as *Assessing Writing*, *The Asia Pacific Education Researcher*, and *Studies in Second Language Learning and Teaching*.