

Students' engagement with ChatGPT feedback: implications for student feedback literacy in the context of generative artificial intelligence

Ying Zhan & Zi Yan

To cite this article: Ying Zhan & Zi Yan (05 Mar 2025): Students' engagement with ChatGPT feedback: implications for student feedback literacy in the context of generative artificial intelligence, Assessment & Evaluation in Higher Education, DOI: [10.1080/02602938.2025.2471821](https://doi.org/10.1080/02602938.2025.2471821)

To link to this article: <https://doi.org/10.1080/02602938.2025.2471821>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 05 Mar 2025.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Students' engagement with ChatGPT feedback: implications for student feedback literacy in the context of generative artificial intelligence

Ying Zhan  and Zi Yan 

Department of Curriculum and Instruction & Assessment Research Centre, The Education University of Hong Kong, Tai Po, N.T, Hong Kong

ABSTRACT



Student feedback engagement is often reported as unsatisfactorily low, undermining the effectiveness of feedback in higher education. The emergence of Generative Artificial Intelligence (GenAI) has created new opportunities to enhance student feedback engagement through its unique affordances. However, assuming that GenAI is a panacea for feedback disengagement can be misleading, as it also presents substantial challenges. Therefore, it is essential to understand how students engage comprehensively with GenAI feedback. However, such understanding is lacking in literature. To close the research gaps, this study investigated how college students engaged with the feedback provided by ChatGPT on their International English Language Testing System writing tasks. Sixteen undergraduates voluntarily participated in unobtrusive observations and stimulated recall interviews. Their feedback engagement was classified into cognitive, metacognitive, affective and behavioural aspects. The findings reveal that participants employed more cognitive strategies than metacognitive ones, experienced minimal emotional resistance, and exhibited varying levels of trust in ChatGPT's feedback. Additionally, the participants' behavioural engagement appeared superficial when requesting feedback, interacting with ChatGPT and revising their writing tasks. Student feedback literacy, including prompt engineering, evaluative judgement, emotional reflexivity, ethical decision-making and meta-cognitive skills, should be cultivated to foster students' active engagement in a GenAI context.

KEYWORDS

Feedback engagement; generative AI; ChatGPT; feedback literacy

Introduction

Feedback is regarded as a fundamental element in teaching and learning in higher education. It is increasingly acknowledged that without active student engagement, feedback can be ineffective for learning (Winstone et al. 2017). However, student feedback engagement is often reported as unsatisfactorily low (Ali, Ahmed, and Rose 2018). Contextual constraints, such as large class sizes, limited curriculum space and time, and heavy teaching workloads, commonly hinder student engagement with feedback (e.g. Carless 2023; Zhan 2023). Feedback engagement is also

CONTACT Ying Zhan  zhanying@eduhk.hk  Department of Curriculum and Instruction & Assessment Research Centre, The Education University of Hong Kong, D1-1F-17, 10 Lo Ping Road, Tai Po, N.T., Hong Kong
© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

closely related to student feedback literacy, which is the knowledge, dispositions and capacities students need to acquire for effective feedback uptake (Carless and Boud 2018; Chong 2021). Research indicates that students' engagement with feedback tends to be superficial when their feedback literacy is not sufficiently developed (Han and Xu 2021).

The emergence of Generative Artificial Intelligence (GenAI) has introduced new opportunities to enhance student feedback engagement, as it holds the potential to address previously identified contextual constraints. GenAI offers six key affordances: interactivity, anxiety-reduction, communicative authenticity, student-centeredness, repetitive practice, and ubiquity (Jeon, Lee, and Choe 2023; Rawas 2024). These features enable GenAI tools to deliver personalised, specific, interactive, and timely feedback through engaging, adaptable, and human-like conversational interactions (Ali et al. 2023). However, it would be risky for educators and researchers to assume that GenAI is a universal solution for student feedback disengagement. GenAI also poses substantial challenges like overdependence on technology and biased or misleading information in the feedback process (e.g. Thorp 2023; Fan et al. 2025). It remains unclear how students can effectively and meaningfully engage with this advanced technology to utilise feedback in ways that support their learning.

More importantly, when interacting with GenAI, students may exhibit unique engagement patterns that require them to think, feel, and act differently compared to how they respond to feedback from human sources (Bozkurt and Sharma 2024). Student feedback engagement and feedback literacy are deeply interconnected (Molloy, Boud, and Henderson 2020; Chong 2021; Wood 2021). Therefore, students may need distinct feedback literacy to effectively engage with the feedback provided by GenAI tools. Although some researchers have started to investigate student feedback literacy in the context of GenAI (e.g. Bearman and Ajjawi 2023; Bearman et al. 2024), these discussions are largely based on scholars' perspectives rather than on students' actual engagement experiences in natural settings.

To address these research gaps, this study aimed to scrutinize how a group of students engaged with the feedback provided by ChatGPT (a popularly used GenAI tool) on their International English Language Testing System (IELTS) writing tasks. The IELTS writing task was chosen because students often use AI to provide feedback on their writing tasks (Imran and Almusharraf 2023). To address the broader audience in higher education, this article focused on identifying general patterns of student engagement with ChatGPT feedback in terms of cognitive, metacognitive, affective, and behavioural dimensions. The IELTS writing task only provided a task situation in which students participated. Through identifying the patterns of student engagement with ChatGPT feedback, the article shed light on student feedback literacy in the context of GenAI, thus benefiting further explorations in this field.

Literature review

Students' engagement with human feedback

Student feedback engagement is pivotal to realising the effects of feedback on learning (Zhang 2022). It goes beyond merely enacting or accepting feedback from others; it requires students to mindfully and emotionally connect with the feedback they receive (Handley, Price, and Millar 2011). The first systematic examination of student feedback engagement in higher education was conducted by Handley, Price, and Millar (2011). They analysed feedback engagement by considering both 'readiness-to-engage' (i.e. the willingness to invest time and effort in completing written assignments and reviewing feedback) and 'active engagement' (i.e. thoughtful reflection on feedback activities). The widely accepted categorisation of feedback engagement is based on the tripartite model of engagement proposed by Fredricks, Blumenfeld, and Paris (2004), which includes cognitive, affective and behavioural aspects (e.g. Ellis 2010; Shen and Chong 2023). Student feedback engagement encompasses how students interpret and pay attention to

feedback (cognitive engagement), how they emotionally react to the feedback (affective engagement) and how they act on feedback (behavioural engagement).

Student feedback engagement can be categorised as either passive or active. Passive engagement involves receiving feedback without active reflection, communication or dialogue (Price, Handley, and Millar 2011). In contrast, active engagement (Handley, Price, and Millar 2011) or agentic engagement (Winstone et al. 2017) emphasise proactive participation. Active engagement focuses on understanding, reflecting on, and applying feedback, while agentic engagement takes this further by encouraging students to take initiative, self-regulate and advocate for their learning needs. It is argued that the effectiveness of feedback on learning is maximised only when students engage with it actively or agentially (Carless and Boud 2018; Winstone and Carless 2019).

Student feedback engagement is determined by the interplay of contextual and individual factors from an ecological perspective (Chong 2021). Contextual factors such as difficult access to teachers, large class sizes, lack of class time for feedback, the power imbalance between teachers and students, end-of-semester comments, insufficient details in comments, lack of chances for resubmission, and teachers' heavy workloads have been identified as constraints on student feedback engagement in traditional feedback environments (e.g. Winstone and Carless 2019; Zhan 2019, 2023; Winstone and Boud 2022; Carless 2023). Individual factors such as learning motivation, feedback literacy, prior feedback experiences and maturity are also influential in student feedback engagement (Shen and Chong 2023).

Among these individual factors, feedback literacy deserves special attention because effectively engaging with feedback demands specific skills and effort from students (van der Kleij 2020). Student feedback literacy encompasses the knowledge, skills, and attitudes necessary for understanding information and utilising it to enhance their work or learning strategies (Carless and Boud 2018). A student who is proficient in feedback literacy can navigate and address the challenges they face while receiving, perceiving, interpreting, and utilising feedback (van der Kleij 2020).

Students' engagement with GenAI feedback

GenAI offers a transformative approach to overcoming contextual constraints in traditional feedback environments. It has the potential to provide students with instant, accessible, and personalised feedback anytime (Hwang et al. 2024) and reduce social anxieties and fear of judgement, enabling students to seek feedback more freely (Tai and Chen 2024). Additionally, GenAI delivers non-judgemental, user-friendly suggestions and can directly assist in revising work or correcting mistakes when requested (Wongvorachan et al. 2022).

However, GenAI has constraints that could impede student feedback engagement. One significant concern is the potential for fostering metacognitive laziness, where students may become overly dependent on GenAI-generated feedback. This dependency could lead to disengagement from deeper cognitive processing and a lack of effort in internalising feedback and applying it meaningfully to their learning (Fan et al. 2025). Therefore, it is essential to explore how students engage metacognitively with the feedback provided by GenAI tools, an aspect that has not been adequately addressed in the traditional tripartite model of feedback engagement.

This study considered a four-dimensional feedback engagement model in a context of GenAI. Cognitive engagement involves how students interpret and judge GenAI feedback, while metacognitive engagement refers to the process by which students actively reflect on, plan and regulate their use of it. Affective engagement refers to their emotional reactions and their attitudes towards GenAI feedback, and behavioural engagement encompasses students' feedback prompts, interactions with GenAI, and revisions they make.

Students' engagement with GenAI feedback differs from their engagement with human feedback due to the interaction with different interlocutors. GenAI is an on-demand and personalised

feedback tool; therefore, the interaction between GenAI and students largely depends on students' proactivity and agency. GenAI may generate different feedback information depending on students' prompts (Tlili et al. 2023). However, the interaction between students and humans (e.g. teachers and peers) differs. Even without students' prompts, the interaction will still occur if their teachers or peers want to have a feedback dialogue with them. Unlike teachers and peers, GenAI cannot clarify questions unaided (Rudolph, Tan, and Tan 2023). Therefore, students need to iteratively determine if they ask appropriate and relevant prompts and adjust their prompts to get the desired information (Bearman et al. 2024). In addition, students' interaction with GenAI may cause less anxiety than that with humans because they do not have to navigate human relationships in any feedback encounters (Rad, Alipour, and Jafarpour 2024). However, students need to manage their trust level in the GenAI system to enact its feedback (Bearman and Ajjawi 2023) and be aware of the ethics of using it in their learning to uphold academic integrity (Wise et al. 2024). Furthermore, interaction with GenAI can leave digital footprints, which can help students iteratively and cyclically reflect on their feedback process. This is difficult to realise in interactions with humans because such interactions are always difficult to trace back. Therefore, it is reasonable to assume that when engaging with GenAI feedback, students will think, feel and act differently.

To actively engage with the feedback provided by GenAI tools, students need to acquire sufficient feedback literacy. Carless and Boud (2018) proposed a framework of student feedback literacy, including 'Appreciating feedback', 'Making judgements', 'Managing affect', and 'Taking action'. Molloy, Boud, and Henderson (2020, 529) proposed a learner-centred framework of feedback literacy which included seven elements, namely, 'Commits to feedback as improvement', 'Appreciates feedback as an active process', 'Elicits information to improve learning', 'Processes feedback information', 'Acknowledges and works with emotions', 'Acknowledges feedback as a reciprocal process' and 'Enacts outcomes of processing of feedback information'. Although these frameworks well characterise the general features of feedback-literate students, they need to be adapted to specific contexts. For instance, Malecka et al. (2022) discovered that the sudden transition to online teaching due to the COVID-19 pandemic introduced an additional contextual factor to feedback practices, influencing the development of student feedback literacy. They suggested that future studies could investigate how students' experiences of interacting with feedback tools and materials shape their feedback literacy. Therefore, from the lens of student engagement with feedback, this study could provide insights into the specificity of student feedback literacy in the context of GenAI.

Methodology

Background

This study was conducted at a teacher education university in Hong Kong, chosen specifically for its conducive environment for integrating GenAI tools into students' learning. The university has recently initiated a pilot program, starting in September 2023, to introduce ChatGPT 3.5 to students and staff. As a result, students have convenient and unrestricted access to ChatGPT 3.5, allowing them to utilise its capabilities and potential in their learning journey fully. The university offers a 5-year Bachelor of Education program with a language policy requiring minimum IELTS scores for graduation (6.0 for most majors). Students must meet these Language Exit Requirements independently, as the university does not provide IELTS preparation courses.

Participants

The participants were recruited using a convenient sampling strategy. A class of students who participated in an elective educational course was approached *via* invitation letters. Sixteen students consented to participate in the study voluntarily. This range of participants is

considered adequate to generate sufficient data to achieve data saturation (Creswell and Poth 2018). They varied in age, gender, year of study and major, and they had not taken the IELTS test before. The demographic information of all the participants is listed in Table 1. The variety of participants resulted in a wealth of data that addressed the research inquiries.

Data collection

Unobtrusive observations and stimulated recall interviews were used in this study to collect data. Unobtrusive observations are often used to capture the interactions among humans or between humans and technology, along with the trace or usage data automatically generated by digital platforms (Jensen et al. 2022). The participants were required to screen-record the whole process when they used ChatGPT to gain feedback on their IELTS writing tasks. A screen capture software – KK Camera, was installed on each participant's laptop for data collection. Before starting unobtrusive observation, each participant was informed of the purpose of using the KK Camera. They were trained to use the software to record their feedback practices with ChatGPT and familiarise themselves with ChatGPT functions. After piloting, the participants chose one of the provided IELTS writing tasks II (see Appendix 1) and wrote about it for 40 minutes. Then, they used ChatGPT to obtain feedback to revise their writing independently. The time spent by them varied from 23 minutes to one hour and 17 minutes.

Stimulated recall interview is a method of introspection where individuals typically watch recorded segments of their behaviour to trigger memories of their thoughts and mental processes they were experiencing at the time (Lyle 2003). After viewing the video records of unobtrusive observations, the first author identified the most critical and typical episodes of students' feedback behaviour with ChatGPT (such as their prompts and revision behaviour as well as their dialogues with ChatGPT) and prepared a personalised semi-structured interview protocol to understand the participants' feedback behaviour, the underlying cognitive and metacognitive process as well as emotional reactions towards the received ChatGPT feedback, which was not directly observed. The participants attended the stimulated recalls a day after they finished their IELTS writing tasks *via* Zoom, which enabled the first author to share the chosen episodes with them. The stimulated recalls were conducted in the participants' mother tongue, lasting about one hour on average.

Data analysis

Text-based interaction between the participants and ChatGPT and their revisions were transcribed for content analysis. These data revealed the participants' behavioural engagement, which was

Table 1. Demographic information of participants in the study.

Code	Age	Gender	Year of study	Major
1	24	Male	4	Math education
2	23	Male	3	General studies education
3	21	Female	3	General studies education
4	20	Female	3	General studies education
5	22	Male	3	AI and IT in education
6	21	Male	4	Kindergarten education
7	23	Female	4	Kindergarten education
8	21	Female	3	General studies education
9	20	Female	3	General studies education
10	23	Female	4	Math education
11	22	Female	2	Math education
12	23	Male	4	Math education
13	20	Female	3	General studies education
14	23	Female	3	General studies education
15	20	Female	4	Math education
16	22	Female	4	Kindergarten education

visible to outsiders through overt actions such as asking questions about the work assessed, interacting with feedback providers, and making revisions (Handley, Price, and Millar 2011). The content of prompts was openly coded and then categorized into eight types. The dialogues between participants and ChatGPT were coded by identifying the number of rounds of interaction. Each round consists of one participant initiating a prompt and ChatGPT responding. If the participant followed up ChatGPT's response, which elicited another ChatGPT's response, this was coded in another round. Three types of revision operations, namely copy, adaptation and rejection (Zhang et al. 2021) were coded.

All the stimulated recall interviews were transcribed for coding. The coding of these data went through the procedures of open coding, axial coding, and selective coding (Williams and Moser 2019). Initial coding was first conducted, and then a category system of student feedback engagement was developed *via* axial coding. A coding matrix of student feedback engagement was established according to cognitive, metacognitive, affective and behavioural engagement aspects discussed in the literature review. In the last step of selective coding, the category system was checked and refined within and across the data sets of participants by repeatedly moving among the raw data, the temporary categories, the research questions, and the literature on student feedback engagement.

Findings

Students' cognitive engagement with ChatGPT feedback

The interview data revealed that the participants majorly used three cognitive strategies to interpret and judge the ChatGPT feedback. These strategies included selective attention, extracting key information and comparison. Eight participants mentioned that they paid selective attention to the specific part of feedback provided by ChatGPT based on self-assessment of their writing weaknesses. The following extract illustrates this point of view.

Researcher: The ChatGPT pointed out two major weaknesses. One was about vocabulary and grammar, and the other was about examples supporting your arguments. Why did you just ask follow-up questions concerning vocabulary and grammar?

Participant #2: The feedback provided by ChatGPT was too long. I needed to focus on the areas that I am not good at. I am not sure if the use of vocabulary is grammatically correct, and I do not know how to use academic-sounding words to enhance my article. So, I would like to know more about this and paid more attention to the relevant information provided by ChatGPT.

Four students talked about using strategies to extract key information provided by ChatGPT. They found it might be challenging to do focused revision later without such cognitive processing since ChatGPT always produced lengthy comments. For example, Participant #1 said:

ChatGPT provided many comments, which sometimes overwhelmed me. I needed to find the key points. My strategy was to read each comment's subtitles or topic sentences, and then I filtered out irrelevant or less significant information.

Ten participants mentioned that they used a comparison strategy to help them judge the usefulness and trustworthiness of ChatGPT feedback. They compared the original version of their writing with the revised version, the comments provided by ChatGPT at different periods, ChatGPT comments with IELTS criteria, and ChatGPT comments with their teachers' previous comments. For example, Participant #16 mentioned that ChatGPT required her to divide her writing passage into more than four paragraphs, and she doubted this comment since her tutor told her that a four-paragraph structure was appropriate for IELTS writing task II. For another instance, Participant #7 asked ChatGPT to score the original version and the revised version and see if the scores

increased and if the revised version addressed the weaknesses mentioned by ChatGPT and improved as expected.

Students' metacognitive engagement with ChatGPT feedback

Compared with cognitive engagement, the participants' metacognitive engagement seemed less. Only five participants reported that they monitored their revision process. For example, Participant #11 mentioned that after she revised one paragraph, she would go back to the ChatGPT feedback to double-check if she really addressed the writing problems identified by ChatGPT. For another instance, Participant #12 proofread the revised paragraph and checked the appropriateness of using the vocabulary suggested by ChatGPT in her sentences.

Six participants reflected on their feedback process with ChatGPT after they finished the revision. They mentioned that they went through all the comments provided by ChatGPT and thought about the usefulness of the ChatGPT feedback in improving their English writing quality. The following extract illustrates this point of view.

Researcher: After ChatGPT provided revisions, you returned to your original version and your previous dialogue with ChatGPT. Why is that?

Participant #14: It made revisions for me, and I understood what improvements and suggestions it provided. At that time, I was thinking whether I could make those revisions on my own. I was thinking about how to fully address the topic, how much to write, and things like that.

Students' affective engagement with ChatGPT feedback

Students' emotional reactions to the ChatGPT feedback were described by the majority of participants ($n=13$) as 'calm', 'relaxed', 'stress-free', and 'unstrained' in most cases when interacting with ChatGPT. For example, Participant #7 described her emotional reactions as follows.

Interacting with ChatGPT is easier because there is no need to worry about asking the wrong questions. However, when facing a teacher, there may be concerns like whether asking such long questions would appear foolish. There are fewer worries. ChatGPT is not my teacher or peer, so even if it points out some weaknesses in my writing, I could calmly face it and think about it.

However, the participants' affective engagement was not always positive; it was sometimes negative. Eight students mentioned that they occasionally felt impatient, confused and upset when ChatGPT misunderstood their prompts, gave irrelevant information and continued to give similar comments even if they had revised accordingly. For example, Participant #2 mentioned that:

I felt bored and impatient when ChatGPT continuously gave me similar suggestions on sentence structure and vocabulary. I was confused about its suggestion since I had already used the suggested vocabulary and did not know why it kept asking me to do the same revision.

Another interesting finding concerns the participants' trust or distrust of ChatGPT. The participants appeared to trust ChatGPT's suggestions on English language forms ($n=11$) but doubted the examples and evidence provided by ChatGPT ($n=12$). For example, Participant #13 claimed, 'I think ChatGPT is better at checking language mistakes than me. So, I basically accepted the language revision provided by ChatGPT'. In contrast, Participant #11 said, 'In terms of content, it can be somewhat general and unconvincing'. Some participants also said that they doubted the usefulness of ChatGPT feedback in improving their IELTS writing performance. They were concerned that ChatGPT was not a real IELTS examiner and might not judge their work accurately and provide helpful feedback for improvement as the teacher did.

Students' behavioural engagement with ChatGPT feedback

Feedback prompts made by the participants

In total, the participants proposed 308 prompts. The least number of prompts provided by individual participants was 9, while the most was 40. The prompts were categorised into eight types according to their content. The number and percentage of each type of prompt are summarized in Table 2, and the example of each prompt type is listed in Appendix 2. Compared with the participants' feedback prompts in global aspects of IELTS writing ($n=55$), such as writing content and structure, more prompts focused on the local aspects ($n=104$), such as grammar, vocabulary and sentence structures. Another interesting finding is that 11.3% of the prompts were a direct copy of the participant's writing passage. In most cases, such prompts occurred at the beginning of the feedback process. The percentage breakdown for requests related to direct revision, general comments, IELTS grades, IELTS criteria or high-mark writing samples was 13.0%, 10.7%, 8.8%, and 2.9%, respectively. Only 1.6% of prompts were related to verifying the credibility of ChatGPT feedback.

Interactions between the participants and ChatGPT in the feedback process

Analysis of participants' dialogues with ChatGPT showed that most observed interaction was one-off ($n=133$). Figure 1 is a screen-captured example of the one-off interaction between a participant and ChatGPT. In addition, 34 two-round interactions with ChatGPT and 18 three-round interactions were recorded. There were 11 more than three-round interactions between the participants and ChatGPT.

Revision operations with the ChatGPT feedback

In total, 308 prompts were made by the participants, which elicited 268 valid comments from the ChatGPT which was relevant to the prompts. These valid comments gave the participants suggestions for revision or direct revision. It reveals that 56.3% of the ChatGPT comments were honestly followed, 27.6% of them were adapted, and 16.1% of them were rejected by the participants.

By closely examining the comments provided by ChatGPT, which the participants honestly followed, most were related to local aspects of IELTS writing, such as grammar, vocabulary and sentence structures or provided direct revision. On the contrary, some participants seemed afraid to directly use the examples and content provided by ChatGPT in their revision. It was influenced by their beliefs of academic misconduct. For example, Participant #6 said:

Actually, when writing academic articles, one should not directly copy from others. Even if there is such technology, the article is written by you, not by AI. You should make your modifications. Only in this way can you learn how to make changes.

Table 2. Prompts made by the participants.

Content of prompts	Number of prompts	Percentage of total prompts
1. Local aspects of IELTS writing	104	33.8%
2. Global aspects of IELTS writing	55	17.9%
3. Direct revision	40	13.0%
4. Direct copying of IELTS writing draft	35	11.3%
5. General requests	33	10.7%
6. IELTS grades	27	8.8%
7. IELTS criteria or high-mark writing samples	9	2.9%
8. Credibility of ChatGPT feedback	5	1.6%
Total	308	100%

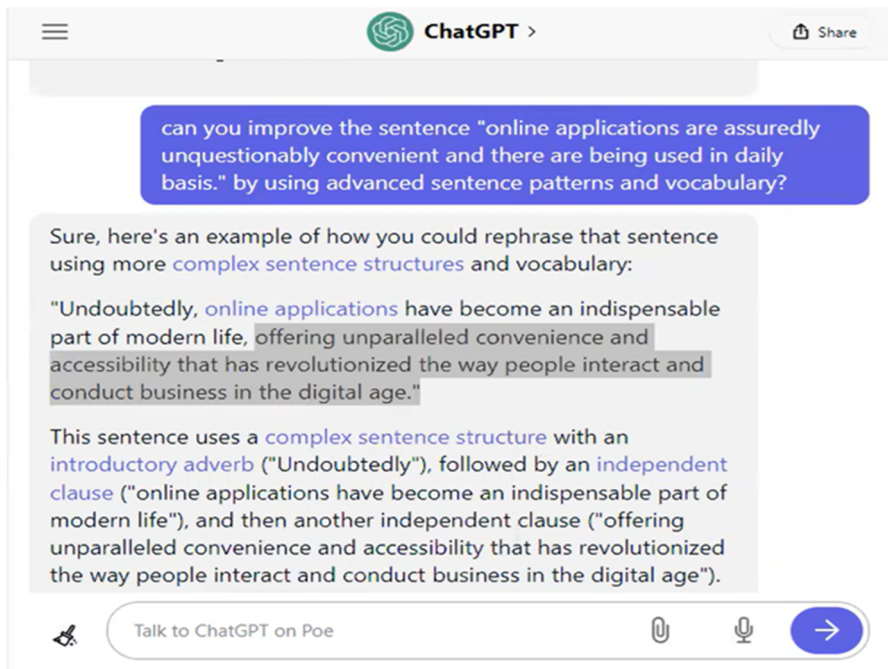


Figure 1. A screen-captured example of the one-turn interaction between a participant and ChatGPT.

Discussion and conclusion

Students' cognitive engagement and evaluative judgement

In this study, the participants mainly adopted three cognitive strategies: selective attention, extracting key information, and comparison. This could be regarded as active engagement (Handley, Price, and Millar 2011). In previous studies, researchers identified that students used similar cognitive strategies, such as noticing and understanding (e.g. Zhang 2017; Koltovskaia 2020; Ranalli 2021). More importantly, this study identified the participants' use of comparison to process the information provided by ChatGPT. Nicol (2021) describes comparison as a hub for the interaction between external information and internal feedback and suggests that explicit comparisons and multiple sequential comparisons generate more feedback benefits. In the context of GenAI, ongoing and natural feedback comparisons with various information sources become more explicit and integrating multiple feedback sources over time becomes more achievable. The participants compared ChatGPT's comments with their teachers' previous feedback and their self-assessment against IELTS criteria to make a judgement. Additionally, they instantaneously compared the original version of their writing with the revised version, as well as the feedback provided by ChatGPT over time, which might not be easily done in a traditional feedback environment.

Using comparison strategies means that students continuously exercise their evaluative judgement (Bearman et al. 2024). Evaluative judgement is the key feature of student feedback literacy and is especially needed in the context of GenAI because GenAI information can appear plausible and relevant even in hallucinations (Bearman et al. 2024). However, six participants did not use comparison strategies, leading them to engage with ChatGPT feedback superficially. Therefore, this study implies that a feedback-literate student in the context of GenAI should acquire sophisticated skills which involve abilities to compare and evaluate different sources of feedback including GenAI feedback and human feedback, assess their work against established criteria, and discern the change of GenAI feedback over time.

Students' metacognitive engagement and metacognitive strategies

Compared with cognitive engagement, fewer participants experienced metacognitive engagement with ChatGPT. In this study, most participants tended to simply read the ChatGPT feedback and make judgements but seldom used it for self-regulation or planning for future work. The participants did not take the initiative to monitor their feedback process with ChatGPT, which caused passive engagement (Price, Handley, and Millar 2011). This echoes metacognitive laziness reported by Fan et al. (2025). Fan et al. (2025) conducted a randomised experiment that indicated ChatGPT might encourage students to rely more on technology, potentially leading to metacognitive laziness, which impeded their ability to self-regulate and engage deeply in the learning process. They also found that ChatGPT feedback could significantly enhance short-term task performance, but it might not improve knowledge transfer. This study provides qualitative evidence to demonstrate that 'metacognitive laziness may prompt short-term performance improvements and long-term skill stagnation' (Fan et al. 2025, 19). By examining the participants' revision operations, it was found that most revisions focused on the local aspects of L2 writing, such as grammar, sentence structure, and vocabulary. This focus may not effectively promote the development of writing skills that can be transferred to other writing tasks (Zhang 2020).

To prevent metacognitive laziness in the context of GenAI, students need to develop metacognitive strategies, which are also emphasised as the ability to enact feedback in the feedback literacy framework proposed by Molloy, Boud, and Henderson (2020). Metacognitive strategies are extremely important for feedback-literate students in the context of GenAI since GenAI tends to provide students with instant, detailed and persuasive information (Yang and Zhang 2024) that makes students experience less cognitive difficulty to trigger their more profound metacognitive process (Alter et al. 2007). In this study, only one-third of the participants used monitoring and reflection strategies to adjust their engagement with the ChatGPT feedback. None of them used strategies of goal-setting and planning. Therefore, to be feedback-literate, students must set feedback goals, plan the subsequent prompts to elicit the desired feedback, self-observe and self-monitor the interaction with GenAI and reflect on the whole feedback process.

Students' affective engagement and emotional reflexivity

In this study, the participants did not report much about emotional resistance, which has been reported in traditional feedback environments (e.g. Winstone et al. 2017; Zhan 2019, 2024). They appeared comfortable and relaxed when interacting with ChatGPT in the feedback process, which indicates positive affective engagement. This echoes the findings of Tai and Chen (2024), who found that students exhibited reduced anxiety when interacting with Chatbots.

Although they do not need to manage negative emotions as described in the previous feedback literacy frameworks (Carless and Boud 2018; Molloy, Boud, and Henderson 2020), students are required to have the ability to do emotional reflexivity which is closely related to their trust within GenAI systems (Bearman and Ajjawi 2023). The trust level in GenAI systems to provide feedback might differ among individuals. According to Ranalli's (2021) and Koltovskaia's (2020) studies, some students consistently doubt AI feedback, whereas others unquestioningly rely on it. In our study, such variations were also observed. Most participants expressed their trust in ChatGPT's feedback on language errors but doubted its suggestions on the content or structure of writing. Some participants also doubted the legitimacy of ChatGPT as an IELTS examiner. Bearman and Ajjawi (2023, 1169) believed that 'highlighting the need to be emotionally reflexive can prompt students to examine their interactions with AI systems in a critical way.' Emotional reflexivity is a process of acknowledging and understanding how emotions influence action (Olson et al. 2021). Therefore, a feedback-literate student in the context of GenAI needs the ability to balance their trust and doubt by understanding GenAI's capabilities and limitations.

Students' behavioural engagement, prompt engineering and ethical decision-making

The participants' behavioural engagement appeared superficial in requesting feedback, interacting with ChatGPT and revising their writing tasks. The study found that 22% of the prompts were general or just copying the writing passages, and 13% of prompts explicitly requested direct revision. It also revealed that most of the interaction between the participants and ChatGPT was one-off. Furthermore, the participants demonstrated a high uptake rate of ChatGPT feedback at 83.9%, with 56.3% of revisions honestly following the suggestions provided by ChatGPT.

Prompt engineering is necessary for a feedback-literate student to seek feedback from GenAI tools to prevent superficial behavioural engagement. Molloy, Boud, and Henderson (2020) discuss seeking skills in traditional feedback contexts. In the context of GenAI, seeking skills are closely related to prompt engineering, which is the process of designing effective questions or stimuli, known as 'prompts', for large language models. The quality of the information provided by GenAI is largely determined by the clarity and accuracy of prompts (Lo 2023). Lo (2023) developed the CLEAR framework, which comprises five key elements of effective prompts: concise, logical, explicit, adaptive and reflective. Students are expected to write accurate, well-structured and specific prompts. Meanwhile, they can explore alternative prompts and continuously evaluate and improve them.

Ethical decision-making is equally vital when using GenAI, as it ensures academic integrity (Wise et al. 2024). Academic integrity requires students to produce work that reflects their understanding and effort while appropriately acknowledging the contributions of others, including GenAI systems. In this study, the high uptake rate of ChatGPT feedback and the significant proportion of revisions following its suggestions highlight the potential for over-reliance on GenAI-generated content. While GenAI can provide valuable support, ethical decision-making requires users to critically evaluate and contextualise GenAI feedback rather than accepting it directly without reflection. Therefore, students should be able to make ethical decisions about how, when and why to use GenAI feedback to ensure that their work remains authentic, credible, and a true representation of their intellectual capabilities.

Figure 2 synthesises the above discussion on student feedback engagement and its implications for student feedback literacy in a GenAI context. The relationship between these elements is bidirectional and mutually reinforcing (Molloy, Boud, and Henderson 2020; Chong 2021; Wood 2021). The study expands the traditional tripartite model of feedback engagement into a

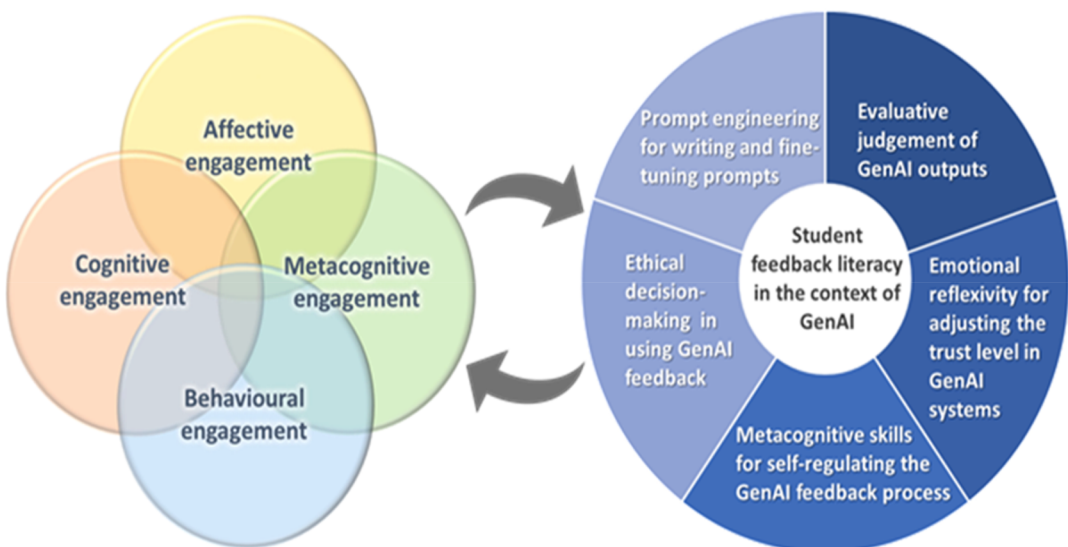


Figure 2. Interplay between student feedback engagement and feedback literacy in a GenAI context.

four-dimensional model that emphasises metacognitive engagement in the GenAI context, addressing the potential risk of metacognitive laziness. Additionally, it proposes five specific aspects of student feedback literacy in the GenAI context that merit further exploration in the field. By documenting students' initial experiences with GenAI tools in the feedback process, this study lays the groundwork for future research on student feedback engagement and feedback literacy in an increasingly GenAI-driven world.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by virtual interdisciplinary research centre fund (No.: 04A47) in the Education University of Hong Kong.

Notes on contributors

Dr Ying Zhan is an Associate Professor at the Department of Curriculum and Instruction and Assessment Research Centre, The Education University of Hong Kong. Her research focuses on assessment for learning, feedback and e-assessment.

Prof. Zi Yan is a Professor at the Department of Curriculum and Instruction and Assessment Research Centre, The Education University of Hong Kong. His research interests focus on two related areas, i.e., educational assessment in the school and higher education contexts; and Rasch measurement, in particular its application in educational and psychological research.

ORCID

Ying Zhan  <http://orcid.org/0000-0002-2362-2887>

Zi Yan  <http://orcid.org/0000-0001-9305-884X>

References

- Ali, F., D. Choy, S. Divaharan, H. Y. Tay, and W. Chen. 2023. "Supporting Self-Directed Learning and Self-Assessment Using TeacherGAI, a Generative AI Chatbot Application: Learning Approaches and Prompt Engineering." *Learning: Research and Practice* 9 (2): 135–147. doi:10.1080/23735082.2023.2258886.
- Ali, N., L. Ahmed, and S. Rose. 2018. "Identifying Predictors of Students' Perception of and Engagement with Assessment Feedback." *Active Learning in Higher Education* 19 (3): 239–251. doi:10.1177/1469787417735609.
- Alter, A. L., D. M. Oppenheimer, N. Epley, and R. N. Eyre. 2007. "Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning." *Journal of Experimental Psychology. General* 136 (4): 569–576. doi:10.1037/0096-3445.136.4.569.
- Bearman, M., and R. Ajjawi. 2023. "Learning to Work with the Black Box: Pedagogy for a World with Artificial Intelligence." *British Journal of Educational Technology* 54 (5): 1160–1173. doi:10.1111/bjet.13337.
- Bearman, M., J. Tai, P. Dawson, D. Boud, and R. Ajjawi. 2024. "Developing Evaluative Judgement for a Time of Generative Artificial Intelligence." *Assessment & Evaluation in Higher Education* 49 (6): 893–905. doi:10.1080/02602938.2024.2335321.
- Bozkurt, A., and R. Sharma. 2024. "Are we Facing an Algorithmic Renaissance or Apocalypse? Generative AI, Chatbots, and Emerging Human-Machine Interaction in the Educational Landscape." *Asian Journal of Distance Education* 19 (1): 1–12. doi:10.5281/ZENODO.10791959.
- Carless, D. 2023. "Teacher Feedback Literacy, Feedback Regimes and Iterative Change: Towards Enhanced Value in Feedback Processes." *Higher Education Research & Development* 42 (8): 1890–1904. doi:10.1080/07294360.2023.2203472.
- Carless, D., and D. Boud. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. doi:10.1080/02602938.2018.1463354.
- Chong, S. W. 2021. "Reconsidering Student Feedback Literacy from an Ecological Perspective." *Assessment & Evaluation in Higher Education* 46 (1): 92–104. doi:10.1080/02602938.2020.1730765.
- Creswell, J. W., and C. N. Poth. 2018. *Qualitative Inquiry & Research Design: Choosing among Five Approaches*. 4th ed. Los Angeles: Sage.

- Ellis, R. 2010. "Epilogue: A Framework for Investigating Oral and Written Corrective Feedback." *Studies in Second Language Acquisition* 32 (2): 335–349. doi:10.1017/S0272263109990544.
- Fan, Y., L. Tang, H. Le, K. Shen, S. Tan, Y. Zhao, Y. Shen, X. Li, and D. Gašević. 2025. "Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance." *British Journal of Educational Technology* 56 (2): 489–530. doi:10.1111/bjet.13544.
- Fredricks, J. A., P. C. Blumenfeld, and A. H. Paris. 2004. "School Engagement: Potential of the Concept, State of the Evidence." *Review of Educational Research* 74 (1): 59–109. doi:10.3102/00346543074001059.
- Han, Y., and Y. Xu. 2021. "Student Feedback Literacy and Engagement with Feedback: A Case Study of Chinese Undergraduate Students." *Teaching in Higher Education* 26 (2): 181–196. doi:10.1080/13562517.2019.1648410.
- Hwang, W. Y., B. C. Guo, A. Hoang, C. C. Chang, and N. T. Wu. 2024. "Facilitating Authentic Contextual EFL Speaking and Conversation with Smart Mechanisms and Investigating Its Influence on Learning Achievements." *Computer Assisted Language Learning* 37 (7): 1632–1658. doi:10.1080/09588221.2022.2095406.
- Imran, M., and N. Almusharaf. 2023. "Analyzing the Role of ChatGPT as a Writing Assistant at Higher Education Level: A Systematic Review of the Literature." *Contemporary Educational Technology* 15 (4): ep464. doi:10.30935/cedtech/13605.
- Handley, K., M. Price, and J. Millar. 2011. "Beyond 'Doing Time': Investigating the Concept of Student Engagement with Feedback." *Oxford Review of Education* 37 (4): 543–560. doi:10.1080/03054985.2011.604951.
- Jensen, L. X., M. Bearman, D. Boud, and F. Konradsen. 2022. "Digital Ethnography in Higher Education Teaching and Learning—a Methodological Review." *Higher Education* 84 (5): 1163–1164. doi:10.1007/s10734-022-00869-x.
- Jeon, J., S. Lee, and H. Choe. 2023. "Beyond ChatGPT: A Conceptual Framework and Systematic Review of Speech-Recognition Chatbots for Language Learning." *Computers & Education* 206: 104898. doi:10.1016/j.compedu.2023.104898.
- Koltovskaia, S. 2020. "Student Engagement with Automated Written Corrective Feedback (AWCF) Provided by Grammarly: A Multiple Case Study." *Assessing Writing* 44: 100450. doi:10.1016/j.asw.2020.100450.
- Lyle, J. 2003. "Stimulated Recall: A Report on Its Use in Naturalistic Research." *British Educational Research Journal* 29 (6): 861–878. doi:10.1080/0141192032000137349.
- Lo, L. S. 2023. "The CLEAR Path: A Framework for Enhancing Information Literacy through Prompt Engineering." *The Journal of Academic Librarianship* 49 (4): 102720. doi:10.1016/j.jacalib.2023.102720.
- Malecka, B., D. Boud, J. Tai, and R. Ajjawi. 2022. "Navigating Feedback Practices across Learning Contexts: Implications for Feedback Literacy." *Assessment & Evaluation in Higher Education* 47 (8): 1330–1344. doi:10.1080/02602938.2022.2041544.
- Molloy, E., D. Boud, and M. Henderson. 2020. "Developing a Learning-Centred Framework for Feedback Literacy." *Assessment & Evaluation in Higher Education* 45 (4): 527–540. doi:10.1080/02602938.2019.1667955.
- Nicol, D. 2021. "The Power of Internal Feedback: Exploiting Natural Comparison Processes." *Assessment & Evaluation in Higher Education* 46 (5): 756–778. doi:10.1080/02602938.2020.1823314.
- Olson, R. E., A. Smith, P. Good, E. Neate, C. Hughes, and J. Hardy. 2021. "Emotionally Reflexive Labour in End-of-Life Communication." *Social Science & Medicine* (1982) 291: 112928. doi:10.1016/j.socscimed.2020.112928.
- Price, M., K. Handley, and J. Millar. 2011. "Feedback: Focusing Attention on Engagement." *Studies in Higher Education* 36 (8): 879–896. doi:10.1080/03075079.2010.483513.
- Rad, H. S., R. Alipour, and A. Jafarpour. 2024. "Using Artificial Intelligence to Foster Students' Writing Feedback Literacy, Engagement, and Outcome: A Case of Wordtune Application." *Interactive Learning Environments* 32 (9): 5020–5040. doi:10.1080/10494820.2023.2208170.
- Ranalli, J. 2021. "L2 Student Engagement with Automated Feedback on Writing: Potential for Learning and Issues of Trust." *Journal of Second Language Writing* 52: 100816. doi:10.1016/j.jslw.2021.100816.
- Rawas, S. 2024. "ChatGPT: Empowering Lifelong Learning in the Digital Age of Higher Education." *Education and Information Technologies* 29 (6): 6895–6908. doi:10.1007/s10639-023-12114-8.
- Rudolph, J., S. Tan, and S. Tan. 2023. "ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?" *Journal of Applied Learning & Teaching* 6 (1):342–363. doi:10.37074/jalt.2023.6.1.9.
- Shen, R., and S. W. Chong. 2023. "Learner Engagement with Written Corrective Feedback in ESL and EFL Contexts: A Qualitative Research Synthesis Using a Perception-Based Framework." *Assessment & Evaluation in Higher Education* 48 (3): 276–290. doi:10.1080/02602938.2022.2072468.
- Tai, T. Y., and H. H. J. Chen. 2024. "The Impact of Intelligent Personal Assistants on Adolescent EFL Learners' Listening Comprehension." *Computer Assisted Language Learning* 37 (3): 433–460. doi:10.1080/09588221.2022.2040536.
- Thorp, H. H. 2023. "ChatGPT is Fun, but Not an Author." *Science (New York, N.Y.)* 379 (6630): 313–313. doi:10.1126/science.adg7879.
- Tili, A., B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang. 2023. "What If the Devil is my Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education." *Smart Learning Environments* 10 (1): 15. doi:10.1186/s40561-023-00237-x.
- van der Kleij, F. M. 2020. "Evaluation of the 'Feedback Engagement Enhancement Tool'to Examine and Enhance Students' Engagement with Feedback on Their Writing." *Studies in Educational Evaluation* 66: 100907. doi:10.1016/j.stueduc.2020.100907.
- Williams, M., and T. Moser. 2019. "The Art of Coding and Thematic Exploration in Qualitative Research." *International Management Review* 15 (1): 45–55.

- Winstone, N. E., and D. Boud. 2022. "The Need to Disentangle Assessment and Feedback in Higher Education." *Studies in Higher Education* 47 (3): 656–667. doi:10.1080/03075079.2020.1779687.
- Winstone, N., and D. Carless. 2019. *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. New York: Routledge.
- Winstone, N. E., R. A. Nash, M. Parker, and J. Rowntree. 2017. "Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes." *Educational Psychologist* 52 (1): 17–37. doi:10.1080/00461520.2016.1207538.
- Wise, B., L. Emerson, A. Van Luyn, B. Dyson, C. Bjork, and S. E. Thomas. 2024. "A Scholarly Dialogue: Writing Scholarship, Authorship, Academic Integrity and the Challenges of AI." *Higher Education Research & Development* 43 (3): 578–590. doi:10.1080/07294360.2023.2280195.
- Wood, J. 2021. "A Dialogic Technology-Mediated Model of Feedback Uptake and Literacy." *Assessment & Evaluation in Higher Education* 46 (8): 1173–1190. doi:10.1080/02602938.2020.1852174.
- Wongvorachan, T., K. W. Lai, O. Bulut, Y.-S. Tsai, and G. Chen. 2022. "Artificial Intelligence: Transforming the Future of Feedback in Education." *Journal of Applied Testing Technology* 23 (1): 95–116.
- Yang, X., and M. Zhang. 2024. "GenAI Distortion: The Effect of GenAI Fluency and Positive Affect." arXiv Preprint arXiv:2404.17822
- Zhan, Y. 2019. "Conventional or Sustainable? Chinese University Students' Thinking about Feedback Used in Their English Lessons." *Assessment & Evaluation in Higher Education* 44 (7): 973–986. doi:10.1080/02602938.2018.1557105.
- Zhan, Y. 2023. "Beyond Technology: Factors Influencing the Effects of Teachers' Audio Feedback on Students' Project-Based Learning." *Technology, Pedagogy and Education* 32 (1): 91–103. doi:10.1080/1475939X.2022.2093965.
- Zhan, Y. 2024. "Are They Ready? An Investigation of University Students' Difficulties in Peer Assessment from Dual Perspectives." *Teaching in Higher Education* 29 (4): 823–840. doi:10.1080/13562517.2021.2021393.
- Zhang, Z. 2017. "Student Engagement with Computer-Generated Feedback: A Case Study." *ELT Journal* 71 (3): 317–328. doi:10.1093/elt/ccw089.
- Zhang, Z. 2020. "Engaging with Automated Writing Evaluation (AWE) Feedback on L2 Writing: Student Perceptions and Revisions." *Assessing Writing* 43: 100439. doi:10.1016/j.asw.2019.100439.
- Zhang, Z. 2022. "Promoting Student Engagement with Feedback: Insights from Collaborative Pedagogy and Teacher Feedback." *Assessment & Evaluation in Higher Education* 47 (4): 540–555.
- Zhang, Y., Z. Pi, L. Chen, X. Zhang, and J. Yang. 2021. "Online Peer Assessment Improves Learners' Creativity: Not Only Learners' Roles as an Assessor or Assessee, but Also Their Behavioral Sequence Matter." *Thinking Skills and Creativity* 42: 100950. doi:10.1016/j.tsc.2021.100950.

Appendix 1. Writing tasks chosen by the participants

- Task 1: Online social media is increasingly being used by people from all over the world to communicate. Why is this happening? Is this generally a good thing or a bad thing? You should write at least 250 words.
- Task 2: Recent advancements in technology have led to many changes in our daily lives, but some people would argue that we are becoming too dependent on technology and that not all these changes are beneficial. Give your opinion and support your argument. You should write at least 250 words.
- Task 3: Many countries around the world are losing their cultural identity and are beginning to appear the same. What are the possible reasons for this? Do you think this is a positive or negative development? Give reasons for your answer and include any relevant examples from your own knowledge or experience. You should write at least 250 words.

Appendix 2. Example of each prompt type identified in this study

Content of prompts	Example
1. Local aspects of IELTS writing	As if you were an IELTS examiner, please provide feedback on how I can improve to achieve a score of 7 in grammar diversity and accuracy in my IELTS essay.
2. Global aspects of IELTS writing	Can you give me more information on how to develop the point of globalization further in my essay introduction by providing some points and ideas?
3. Direct revision	Based on the introduction I gave to you, correct the introduction which matches the requirement for IELTS standards
4. Direct copying of IELTS writing draft	Participants just copied their writing in the dialogue box.
5. General requests	Can you give me some advice on my IELTS writing?
6. IELTS grades	What grade can I get in this IELTS writing?
7. IELTS criteria or high-mark writing samples	Write me a 9-mark example of IELTS on the topic of "Recent advancements in technology..."
8. Credibility of ChatGPT feedback	How do you get the high-scoring essays scored by human examiners on previous IELTS exams?