Investigating student engagement with AI-driven feedback in translation revision: A mixed-methods study

Simin Xu[1] · Yanfang Su[1] · Kanglong Liu[1]

Abstract

Despite the well-established importance of feedback in education, the application of Artificial Intelligence (AI)-generated feedback, particularly from language models like ChatGPT, remains understudied in translation education. This study investigates the engagement of Master's students in translation with ChatGPT-generated feedback during their revision process. A mixed-methods approach, combining a translation-and-revision experiment with quantitative and qualitative analyses, was employed to examine the feedback, translations before and after revision, the revision process, and student reflections. The results reveal complex interrelations among cognitive, affective, and behavioural dimensions influencing students' engagement with AI feedback and their subsequent revisions. Specifically, the findings indicate that students invested considerable cognitive effort in the revision process, despite finding the feedback comprehensible. Moreover, they exhibited moderate affective satisfaction with the feedback model. Behaviourally, their actions were largely influenced by cognitive and affective factors, although some inconsistencies were observed. This research provides novel insights into the potential applications of AI-generated feedback in translation teaching and opens avenues for further investigation into the integration of AI tools in language teaching settings.

Keywords Student engagement · AI-generated feedback · Translation education · ChatGPT · Revision process

✉ Kanglong Liu
  kl.liu@polyu.edu.hk

  Simin Xu
  siminxu@polyu.edu.hk

  Yanfang Su
  yanfang.su@connect.polyu.hk

[1] Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China

🖄 Springer

1 Introduction

Feedback is a crucial component of educational practice, with a profound impact on student achievement (Hattie, 1992). In the context of translation teaching, feedback plays a vital role in guiding students toward improving their translation skills (Bruton, 2007). However, the traditional model of feedback, where teachers manually craft responses to student work, is time-consuming and often places a significant strain on teachers (Guo et al., 2024; Shen et al., 2017). The emergence of artificial intelligence (AI) tools, such as ChatGPT, offers a promising solution to enhance teacher efficiency and provide high-quality feedback. Recent studies have explored the potential of ChatGPT to generate personalized, timely, and detailed feedback (AlGhamdi, 2024; Banihashem et al., 2024; Guo & Wang, 2023; Su et al., 2023). However, the majority of these studies have focused on writing, with only a few examining the application of ChatGPT in translation education (Cao & Zhong, 2023). Moreover, research has shown that the effectiveness of feedback is not solely determined by its quality, but also by how students engage with it. This engagement includes three dimensions: cognitive (the depth of processing of feedback), affective (emotional and attitudinal responses to feedback), and behavioral (revision operations undertaken in response to feedback and observable strategies used for improvement) (Handley et al., 2011; Koltovskaia, 2020). However, there is a notable scarcity of research investigating learner engagement with ChatGPT translation feedback. This study seeks to bridge the existing gap by investigating how Master's students majoring in translation interact with ChatGPT translation feedback on cognitive, affective, and behavioural levels. By exploring the potential of ChatGPT in facilitating translation instruction, this research aims to contribute to the advancement of translation pedagogy and inform the effective integration of AI tools in educational settings.

2 Related work

2.1 ChatGPT feedback in language learning and translation

Feedback, a widely employed instructional tool in classrooms, was initially conceptualized as information provided by teachers, peers, books, and texts in response to learners' work or performance (Hattie & Timperley, 2007). In the context of translation, feedback plays a crucial role in supporting and extending learning goals, enabling translators to view their work from the perspectives of readers or users, and fostering the skill of self-assessment (Washbourne, 2014). Previous studies have investigated the efficacy of various types of feedback in translation education. For instance, Yu et al. (2020) demonstrated the effectiveness of written corrective feedback in translation, revealing that students with low L2 proficiency tend to benefit from direct feedback. Similarly, Li and Ke (2022) found that peer feedback not only improved student performance but also enhanced their capacity for evaluative judgments. While the advantages of translation feedback are well-documented, providing

it remains a time-intensive process, and the growing class sizes have further increased teachers' workload (Banihashem et al., 2024; Er et al., 2021). Consequently, teachers face the challenge of balancing the quality and timeliness of feedback when assessing students' work (AlGhamdi, 2024). This dilemma has prompted researchers to explore automated approaches to providing translation feedback that can alleviate the burden on teachers while maintaining the quality of feedback (Han & Lu, 2023).

In this context, the emergence of ChatGPT has introduced a novel perspective on feedback provision. As a generative AI (GenAI) chatbot developed by OpenAI, ChatGPT was trained on a vast corpus of texts (Ekin, 2023). Its capabilities have been highlighted in a range of tasks, including writing and translating (Herbold et al., 2023; Lee, 2023), making it a viable tool in foreign language learning. Several studies have explored the potential of ChatGPT in generating feedback by comparing the characteristics of feedback generated by ChatGPT and teachers. For instance, Steiss et al. (2024) conducted a comparative analysis of the quality of human and ChatGPT feedback on writing assignments, revealing that AI and human feedback exhibited distinct features. Guo and Wang (2023) found that when assessing students' writing, teachers primarily focused on generating content-related and language-related feedback, whereas ChatGPT feedback addressed three aspects (i.e., content, organization, and language) equally. This study also highlighted that teachers held both negative and positive perceptions toward ChatGPT feedback. AlGhamdi (2024) employed a blinded approach to investigate how computing students responded to ChatGPT feedback after using both ChatGPT and human feedback in technical writing. The findings revealed that ChatGPT was capable of generating consistent and detailed feedback. While numerous studies have examined the use of ChatGPT feedback in writing contexts, research exploring the potential of ChatGPT to provide feedback for translation tasks remains relatively underexplored. One notable exception is the study by Cao and Zhong (2023), which examines the effectiveness of feedback generated by ChatGPT and teachers by comparing students' revised translation drafts. However, the study's scope is limited to assessing the quality of the translations, leaving the crucial aspects of students' perceptions of feedback and revision operations unexplored. The ways in which students engage with the revision process also remain unclear. Effective feedback is not solely determined by its content and quality, but also by how it is interpreted and internalized by learners (Nicol & Macfarlane-Dick, 2006). Therefore, to gain a comprehensive understanding of the potential of ChatGPT feedback, it is essential to investigate how students interact with and utilise it during the revision process.

2.2 Learner engagement with feedback

Learner engagement, a multifaceted construct encompassing emotional, cognitive, and behavioural dimensions (Fredricks et al., 2004), is a crucial factor in education. It is widely regarded as a key indicator of the extent to which students are committed to learning (Cheng et al., 2023). This concept is equally relevant to feedback, as its effectiveness is inextricably linked to student engagement (Jørgensen, 2018; Winstone et al., 2017).

To gain a deeper understanding of students' engagement with feedback, researchers have refined its analytical framework (Ellis, 2010; Han & Hyland, 2015; Zhang & Hyland, 2018; Zheng & Yu, 2018; Qian & Li, 2023). Specifically, cognitive engagement refers to the cognitive processes that learners employ in response to feedback (Ellis, 2010). This construct can be further categorised into three sub-components: awareness, cognitive operations, and meta-cognitive operations (Han & Hyland, 2015). Awareness, which is the fundamental level of cognitive engagement, encompasses two key aspects: noticing and understanding. Noticing refers to learners' ability to discern the intention of feedback, while understanding demonstrates the degree to which learners can identify errors and provide accurate explanations. Previous research has measured cognitive operations by examining the macro strategies that learners used to respond to feedback (Pan et al., 2023), as well as the cognitive strategies employed to process feedback and generate revisions. Furthermore, meta-cognitive operations have been identified as actions that regulate mental effort, comprising two dimensions: monitoring and planning (Qian & Li, 2023). Specifically, monitoring is conceptualized as learners' ability to identify additional errors and inaccuracies beyond those highlighted in the feedback. Planning strategies, on the other hand, involve learners' prioritisation when addressing feedback, which help reduce cognitive load. Although directly observing cognitive engagement is difficult, it can be measured indirectly through questionnaires and stimulated recall protocols (Chen, 2021; Philp & Duchesne, 2016). To overcome this challenge, the present study adopts a mixed methods approach, combining questionnaires, interviews, and revision records to capture learners' understanding of feedback, their application of cognitive operations, and their use of metacognitive operations during the revision process.

Affective engagement, also known as emotional engagement, encompasses learners' affective responses to feedback (Ellis, 2010). According to Han and Hyland (2015), this construct is characterized by the emotions experienced upon receiving feedback and revising one's work, as well as attitudinal responses toward feedback. Building on this concept, Zheng and Yu (2018) proposed a framework that distinguishes between three components of affective engagement: affect (learners' emotions and feelings), judgment (positive or negative evaluation of feedback), and appreciation (the perceived value of feedback). To measure affective engagement, researchers commonly employ questionnaires and self-report methods (Fan & Xu, 2020; Guo et al., 2023; Lee et al., 2024; Philp & Duchesne, 2016), which are also adopted in the present study.

Behavioural engagement is closely associated with the actions learners take in response to feedback (Zheng & Yu, 2018). This concept includes revision operations, which pertain to the extent to which learners incorporate feedback, as well as observable strategies employed to enhance their work (Han & Hyland, 2015). For example, Zhang (2017) analysed interview responses and revision time to demonstrate how students behaviourally engage with computer-generated feedback when not under time constraints. Similarly, Tian and Zhou (2020) examined textual changes between learners' initial drafts and revised versions to reflect behavioural engagement. Consistent with these studies, the present research seeks to measure behavioural engagement using three indicators: time spent on revisions, revision operations, and revision strategies.

Building on the aforementioned framework, a growing body of empirical research has explored how students engage with feedback, revealing complex patterns of engagement and diverse ways in which students interact with and respond to feedback. For instance, Zheng and Yu (2018) examined students' engagement with feedback in writing classes and found that engagement was closely tied to language proficiency, resulting in imbalances among the three dimensions of engagement. Similarly, Yu et al. (2019) investigated the engagement of Master's students with peer feedback during second language writing, uncovering a complex relationship both within and across the three dimensions. More recently, Cheng and Zhang (2024) studied how students engaged with AWE (automated writing evaluation)-teacher integrated feedback in writing tasks, finding that students exhibited deeper engagement both behaviourally and cognitively, while displaying similar levels of affective engagement compared to students who received only teacher feedback. Despite the growing body of research, the majority of studies have been conducted in the context of second language writing, with only a few scholars focusing on translation feedback. A notable exception is Zheng et al. (2020), which explored how students engaged with teacher translation feedback. Their study highlighted the interplay among the three dimensions of engagement, revealing their mutual influence and the imbalances that exist among them. However, the study's small sample size (n=3) limits its generalizability, underscoring the need for further research in this area.

3 The present study

As previously discussed, research on student engagement with feedback has primarily concentrated on second language writing, while translation feedback has received limited attention. The recent advent of ChatGPT, an advanced AI tool with sophisticated natural language processing capabilities, offers a promising opportunity to provide feedback on translation assignments. ChatGPT's ability to generate fluent, detailed, and coherent feedback on student assignments in a short time (Dai et al., 2023) can assist teachers in assessing large classes and reducing their workload. However, the effectiveness of ChatGPT in translation teaching remains largely unexplored, and the nature of student engagement with its feedback in specific learning contexts remains unclear.

Drawing on the framework of student engagement with feedback (Zheng & Yu, 2018), this study aims to address the following research question: How do students engage with feedback provided by ChatGPT cognitively, affectively, and behaviourally?

4 Methodology

4.1 Participants

The study was conducted at a university in Hong Kong with a sample of 29 students enrolled in the Master of Translating and Interpreting (MATI) programme. Prior to

the experiment, participants completed an online pre-study survey about their backgrounds. The results revealed that most participants spoke Mandarin Chinese as their first language (L1) (n = 27). Although two students' L1 was Cantonese, all participants were from the Chinese mainland, where Mandarin Chinese is the official language. The sample comprised 21 female students (72.4%) and 8 male students (27.6%), reflecting the gender distribution of the MATI programme. The majority of students (n = 23) had prior experience using ChatGPT (see Table 1). After a detailed explanation of the experimental procedure, all participants provided informed consent by signing a consent form.

4.2 Research procedure and data collection

Our research team developed a specialised AI-powered Translation Teaching Platform that integrated the capabilities of ChatGPT. This platform was specifically designed to support students' translation learning (see Fig. 1). The platform's testing function, which utilised ChatGPT's ability to provide feedback, was employed in this study. A built-in prompt was developed based on the translation assessment rubric from Hurtado Albir & Taylor (2015) to ensure consistent and detailed feedback (see Fig. 2). This prompt was refined multiple times to improve the quality of the feedback. Before data collection, we invited two experienced translation teachers to test the platform and evaluate the feedback. Both confirmed that the feedback quality met their expectations.

The experiment was conducted online via Zoom. Participants were first tasked with translating a Chinese text of approximately 190 words into English within 70 min. Upon completing the translation, the platform generated feedback on their performance using ChatGPT. Participants then revised their translations based on this feedback. To align with classroom teaching resources, we selected a piece of political news as the source text. Two professional translation teachers were consulted to confirm that the text matched the difficulty level of their teaching materials. It is worth noting that, although ChatGPT can generate multiple versions of

Table 1 Demographic information and ChatGPT usage experience of participants	Category	*N*	Percentage
	Gender		
	Male	8	27.6%
	Female	21	72.4%
	First Language		
	Chinese (Mandarin)	27	93.1%
	Chinese (Cantonese)	2	6.9%
	Experience of using ChatGPT		
	Over 10 times	14	48.3%
	5–9 times	2	6.9%
	1–4 times	7	24.1%
	Never	6	20.7%

User Translation Experiment Test TestOld Feedback Personal Home   測试用户1

ChatGPT   **Teacher Feedback**

Peer Feedback   **Machine Translation**
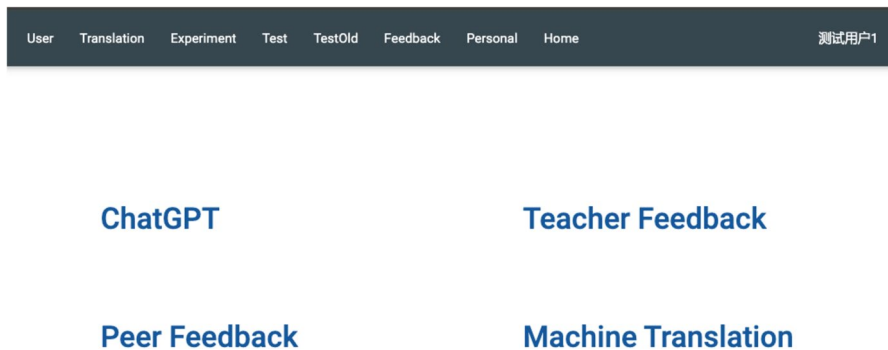
Fig.1 The feedback category of the AI-powered translation teaching platform

Now you are a teacher in translation major. Here is a translation assessment rubric consisting of 3 aspects: Expression of the meaning of the original text accounts for 40% and it concludes same information, same clarity, and same register. Composition in the target language accounts for 40% and it concludes conventions of written language (correct orthography and typography), vocabulary (appropriateness and richness), morphosyntax (good use of verb tenses and modes, prepositions, etc.), cohesion (good use of connectors and referential elements), and coherence (ideas well organized and clearly presented). Level of communication of the target text accounts for 20% and it concludes overall quality of the target text, appropriateness of the genre's conventions, and appropriateness of the translation's purpose and target audience. Please give comments, suggestions and a score out of 100 (including the score of each aspect and the overall score) based on the above mentioned rubric for the following translated text: *translated text*. The source text is: *source text*.

Fig.2 The built-in prompt used to generate ChatGPT feedback

feedback based on the same prompt, only the first output was considered in this study. Furthermore, to ensure that all participants completed their initial translation draft without using machine translation tools or online resources, they were required to share their screens on Zoom during the translation process. The revision process was also recorded on video to analyse participants' behaviour.

Following the revision phase, all participants were required to finish a post-survey which focused on their cognitive and affective engagement with ChatGPT feedback. The survey instrument was adapted from Chen (2021)'s validated questionnaire, which was originally designed to assess students' engagement with peer and teacher feedback. The construct validity and content validity of their questionnaire have been established in their study, with the instrument demonstrating robust

measurement properties for assessing student engagement with feedback. It consists of 15 Likert-scale questions with five response options: 1—strongly disagree, 2—disagree, 3—Neither agree nor disagree, 4—agree, and 5—strongly agree. All the questions in each dimension are divided into different sub-categories (See Tables 4 and 5). Cronbach's alpha values for the survey questions measuring cognitive engagement and affective engagement were 0.807 and 0.802 respectively, demonstrating high internal consistency reliability. To clarify the students' responses of the survey and amplify their revision processes in detail, we conducted follow-up interviews with a subset of participants after the experiment. Following purposeful sampling strategies (Patton, 2002), we selected four participants for in-depth interviews. To maintain the anonymity of the students, we assigned them pseudonyms: Student 1, Student 2, Student 3, and Student 4. We chose them because their survey responses reflected the group's average, ensuring they could provide a representation of the typical trends of all participants. Additionally, their voluntary consent to participate further legitimized their inclusion in the study. Drawing on the framework established by Zheng et al. (2020), the 12 interview questions (see Appendix 1) were designed to elicit (1) reflective commentary on the translation feedback and (2) retrospective accounts of the revision process. All interviews were conducted in Chinese to ensure linguistic and cultural authenticity. To ensure content validity, two researchers reviewed the interview questions before the experiment commenced. Several questions were also refined based on the survey responses of the focal participants. A research assistant collected and collated all relevant data, including students' translation drafts, revision drafts, feedback, post-survey responses, and revision recordings.

4.3 Data analysis

The present study employed a mixed methods approach, incorporating both qualitative and quantitative data. The qualitative data further explained the results of the quantitative analysis. Moreover, this integration provided a comprehensive analysis that captured the complexity of student engagement. The qualitative analysis included text analysis of participants' feedback and drafts, semi-structured interviews, and observational analysis of revision recordings. The study utilised a structured measurement framework to assess learner engagement across three dimensions: cognitive engagement, affective engagement, and behavioural engagement. To facilitate data analysis, the data sources were identified and segmented according to the sub-categories of learner engagement outlined in Table 2, which was refined and adapted from existing engagement frameworks (Zhang & Hyland, 2018; Zheng & Yu, 2018; Qian & Li, 2023).

4.3.1 Analysis of student translations and ChatGPT feedback

One research assistant coded the content of the ChatGPT feedback according to the assessment rubrics designed for the prompt, as well as the framework established by Tian and Zhou (2020). The feedback was categorised into two main

Table 2 Categories and sub-categories of student engagement with relevant sources

Categories	Sub-categories	Relevant sources
Cognitive engagement	Noticing and understanding feedback	Post-survey, interview data
	Meta-cognitive strategies	Post-survey, revision recordings, interview data
	Cognitive operations	Post-survey, revision recordings, interview data
Affective engagement	Emotional responses	Post-survey, interview data
	Attitudinal responses	Post-survey, interview data
Behavioural engagement	Revision operations	Revision recordings, interview data
	Revision strategies	Revision recordings, interview data
	Revision duration	Revision time

types: surface-level and meaning-level feedback. Surface-level feedback refers to comments that do not alter the underlying meaning, including corrections related to written conventions, genre conventions, word tenses, modes, and prepositions. In contrast, meaning-level feedback requires changes to the underlying meaning, including cohesion, coherence, translation accuracy, and lexical choices. To investigate how participants addressed the feedback and revised their translations, we categorised their modifications based on whether they responded to the feedback, and calculated the feedback quantity, feedback uptake quantity, and feedback uptake rate. Furthermore, we observed and analysed all participants' revision recordings to document the time spent on revisions, as well as the type and frequency of strategies employed during the revision process. Three coding examples are presented in Table 3. This comprehensive analysis facilitated a detailed examination of participants' revision behaviours.

4.3.2 Analysis of post-survey and interview data

Regarding the post-survey, we assessed cognitive engagement through items 1–8 and affective engagement through items 9–15. Students' responses to the post-survey provided valuable insights into their experiences during the revision process and their perceptions of the feedback received. The oral interviews were transcribed verbatim using the automatic speech recognition app iFlyRec (https://

Table 3 Examples of students' revision recordings with coding details

Student name	Revision time	Revision strategy	Times
Participant 1	0:51:21	Online searching	6
Participant 2	0:46:21	Online searching	7
		Using corpora	5
		Using dictionary	2
Participant 3	0:54:34	Using dictionary	6
		Online searching	4

www.iflyrec.com) and subsequently proofread manually by a research assistant to ensure accuracy. Two coders thoroughly read the transcripts multiple times to gain a comprehensive understanding of the content, and then conducted coding based on the framework outlined in Table 2. A qualitative analysis was employed to provide a nuanced and in-depth understanding of the interviewees' cognitive, affective, and behavioural engagement. To ensure the trustworthiness of the coding, the two coders cross-referenced students' revision drafts and recordings to validate their responses and resolve any discrepancies. In cases of ambiguity or disagreement, discussions were held to reach a consensus and ensure inter-rater reliability.

5 Findings

Prior to presenting the findings on the three dimensions of engagement, it is essential to illustrate the structure and format of the ChatGPT feedback employed in the current study. Our analysis revealed that the feedback generated by ChatGPT, based on the prompt input, typically consisted of three primary components: (1) a reference translation, (2) comments and suggestions, and (3) a grade (see Appendix 2).

5.1 Cognitive engagement

To investigate how participants engaged with ChatGPT feedback cognitively, we operationalized cognitive engagement into three dimensions: (1) noticing and understanding the feedback, (2) meta-cognitive strategies to monitor the revision process, and (3) cognitive operations in mental activities to recall and utilise feedback. The post-survey items were categorised according to these dimensions (see Table 4). Overall, the mean scores of the eight survey questions ranged from 3.31 to 4.28, indicating that the participants exhibited a high level of cognitive engagement with the ChatGPT feedback. With regard to understanding the feedback, the results of Q1 and Q2 suggested that most participants did not experience significant difficulties in comprehending and revising their work based on the feedback. However, there were also contrasting opinions. Four interviewees noted that while they could understand most suggestions, they found some meaning-level points confusing. For instance, ChatGPT feedback advised Student 1 to focus on coherence and cohesion but failed to specify the exact errors, leaving her uncertain about how to improve the translation.

The meta-cognitive strategies employed by participants in both tasks primarily involved monitoring their mental effort, as well as practicing and planning the revision procedure. As indicated by Q3 and Q4, processing feedback required considerable effort from participants (M = 4.28), and they had to review the feedback repeatedly during the revision process (M = 4.07). This revealed a mismatch between how students understood the feedback and the cognitive strategies they employed. For instance, Student 3 mentioned that she spent a lot of time reading the feedback because she wanted to fully understand it. This suggests that while students were

Table 4 Descriptive statistics for cognitive engagement sub-categories

Sub-categories	Items	M	SD
Noticing and understanding feedback	Q1: I could totally understand the feedback	4.00	0.93
	Q2: After reading the feedback, I could easily make improvement	3.55	1.06
Meta-cognitive strategies	Q3: When I revised my translation, I had to refer to the feedback repeatedly	4.07	0.88
	Q4: I put much effort into revising my translation draft based on the feedback	4.28	0.96
	Q5: I firstly checked the grade and then referred to the feedback	3.76	1.35
	Q6: Compared with other issues, I tended to prioritize surface-level problems	4.10	0.82
Cognitive operations	Q7: When I revised my translation, I totally followed the feedback	3.31	1.07
	Q8: After receiving feedback, I first critically thought about the feedback and then revised my translation draft according to it	4.24	0.79

putting in considerable mental effort to process the feedback, they may not have fully grasped or effectively applied it to improve their work.

Mental effort was also invested in monitoring translation accuracy, which was accompanied by corresponding revision actions. Four interviewees reported actively checking whether their revisions had enhanced the quality of their translations by reviewing their work after completing the revision process. Another example is self-correction. For instance, Student 1 made a revision that was not flagged by ChatGPT:

> *When I read the source text* "布林肯与韩正会晤时毫不讳言，'我们有机会在最近两国高层接触的基础上前进，是一件好事'。"*again during the revision process, I realised that there were quotation marks in this sentence, which I didn't notice when I first translated it. Considering that the source text was news, I thought using direct speech could make the translation more accurate and objective.*

Another meta-cognitive strategy employed by participants was planning. According to Q5, students often prioritised checking their grades when reviewing feedback (M = 3.76). For example, Student 3 stated that she initially reviewed her grade upon receiving the feedback to assess the overall quality of her translation. However, participants exhibited varying preferences when deciding on the priority of addressing feedback (SD = 1.35). During the interview, Student 2 explained that she first consulted reference translations while revising, followed by reviewing the feedback comments, and then checked the grade. She expressed doubts about the reliability of machine-generated grades. The results of Q6 showed that, during the revision process, students tended to focus on correcting surface-level errors before tackling deeper, meaning-related issues (M = 4.10). Student 1 elaborated on her approach:

> *First, I went through the feedback and made corrections as I spotted issues like sentence structure and grammar. Once that was done, I tackled the deeper problems related to word choices and accuracy, which took a bit more effort.*

In terms of cognitive operations, judgment played a key role in processing ChatGPT feedback. The mean score for Q8 (M = 4.24) was higher than that for Q7 (M = 3.31). These results suggest that, although students showed a moderate tendency to follow the feedback completely, they also expressed skepticism about its accuracy and reliability during the revision process. This skepticism prompted corresponding actions, such as making self-initiated changes or choosing not to adopt certain feedback. This inconsistency highlights a conflict within their cognitive operations. Three interviewees reported that they often re-evaluated the accuracy of feedback points that differed from their own translations by consulting additional resources, such as Google and linguistic corpora.

5.2 Affective engagement

According to the conceptual framework in Table 2, affective engagement includes both emotional and attitudinal responses. The analysis of survey and interview data

revealed a complex pattern within this dimension. In terms of emotional responses, students demonstrated positive feelings (see Table 5), aligning with the findings on cognitive engagement. Most students found the revision process enjoyable (M=4.00), and ChatGPT's translation feedback generally boosted their confidence and provided encouragement, motivating them to further refine their translations (M=3.86). A significant factor contributing to this positive affective response was the high grades generated by ChatGPT. Student 1 expressed surprise upon seeing her grade, as she had not anticipated such a high mark for her initial translation. This unexpected result increased her confidence and inspired her to make additional improvements to the translation. However, two students expressed differing perspectives. Student 2 noted that her feelings were not significantly influenced by the grade, explaining that teacher-assigned scores are based on comparing the work of all students, providing a more nuanced understanding of translation proficiency within a classroom context. In contrast, ChatGPT's scoring lacks this comparative insight.

The descriptive statistics in Table 5 showed that while positive feedback significantly boosted students' confidence (M=4.69), negative feedback had little effect on their motivation during the revision process (M=2.93). All the interviewees admitted that they initially felt pleased upon receiving the feedback. However, they later observed that ChatGPT often provided an excessive amount of positive feedback and expressed a preference for more constructive criticism to enhance their translations. Student 1 commented that, at first, she felt quite confident because all the feedback from ChatGPT was positive. Nevertheless, she soon realised that the feedback offered little value in improving her translation skills.

With respect to students' attitudes, the results indicated a positive response to ChatGPT feedback. Overall, students found the content of the feedback engaging (M=3.97) and expressed a strong willingness to continue using it in future translation teaching (M=4.00). The positive remarks gathered from the interviews were largely attributed to students' expectations and enthusiasm for the format of this innovative feedback, as well as the potential of ChatGPT in translation tasks. For example, Students 2 and 4 highlighted that ChatGPT could generate more native-like expressions, while Student 3 observed that the combination of reference translations with specific comments was more useful than receiving only one type of feedback.

Table 5 Descriptive statistics for affective engagement sub-categories

Sub-categories	Items	M	SD
Emotional responses	Q9: I felt confident and encouraged in translation revision after reading the feedback	3.86	1.03
	Q10: I enjoyed the revision process very much	4.00	0.89
	Q11: Positive feedback makes me happy	4.69	0.47
	Q12: Negative feedback makes me frustrated	2.93	1.46
Attitudinal responses	Q13: I'm interested in the content of the ChatGPT feedback	3.97	0.98
	Q14: I expect to receive the ChatGPT feedback	4.00	1.00
	Q15: To what extent do you satisfied with the whole feedback?	3.55	0.78

Furthermore, students demonstrated a moderately favourable response to the effectiveness of ChatGPT feedback (M = 3.55). An analysis of the responses from the four interviewees revealed varying levels of satisfaction with different aspects of the feedback. Specifically, three students expressed notable satisfaction with the feedback on lexical choices, while two students appreciated the suggestions regarding sentence structure. The remarks of Students 2 and 3 effectively capture the opinions of the majority:

> *Student 2: ChatGPT gave me some suggestions for my sentence structure, and I think they could really improve the quality of my translation.*
> *Student 3: I was always worried that my translation was not native. When I looked over ChatGPT's feedback, I noticed that some of the suggested word choices fit much better with the style of news.*

However, they also expressed a certain degree of disappointment. Notably, two students acknowledged that some suggestions were unnecessary. Student 1 articulated her dissatisfaction with ChatGPT feedback:

> *When it comes to the term "分歧", I initially used "division", but ChatGPT suggested changing it to "difference". I didn't think that was necessary.*

Another factor contributing to the students' disappointment was the generic nature of ChatGPT's feedback. All interviewees noted that, although ChatGPT identified issues or offered specific corrections, it would be more helpful if it provided more detailed guidance. This view was shared by Student 3, who remarked during the interview:

> *The feedback noted that my translation lacked clarity, but it provided neither specific areas nor suggestions for improvement.*

Her response suggests that students' disappointment partly stemmed from the confusion they experienced when interpreting ChatGPT's feedback, indicating an alignment between their cognitive and affective engagement.

5.3 Behavioural engagement

Students' behavioural engagement can be analysed through their revision operations, including observable strategies used to improve translation quality and the time spent revising. Feedback uptake and textual modifications reveal how students approached revising their work. As shown in Table 6, the majority of modifications were made in response to ChatGPT's translation feedback (756 items), while a smaller number were self-initiated modifications (96 items). Student 1 recalled that, because ChatGPT did not identify all the issues in her translation, she began to identify them independently to enhance the translation. This also reflected her dissatisfaction with ChatGPT's feedback in the affective domain. Table 7 summarises the content of ChatGPT's feedback and the four participants' revision operations

Table 6 Summary of modifications in translations

Revision type	Amount
Correction based on feedback	756
Self-correction	96
Total	852

Table 7 ChatGPT feedback and learners' uptake rate

Feedback content	Feedback quantity	Feedback uptake quantity	Feedback uptake rate
Surface-level	279	130	47%
Meaning-level	477	304	63.73%
Total	756	434	57.41%

in response. It is evident that ChatGPT predominantly provided suggestions at the meaning level (477 items), achieving an uptake rate of 63.73%. Surface-level feedback accounted for fewer items (279) and had a lower uptake rate of 47%. Although students found surface-level feedback easier to process, they experienced greater difficulty interpreting meaning-level feedback. Nevertheless, the uptake rate for meaning-level feedback was more than twice that of surface-level feedback, highlighting a mismatch between their cognitive and behavioural engagement. This mismatch may be attributed to "the lack of feedback on genre-specific conventions and practical aspects of translation" (Student 3). Furthermore, when reflecting on the revision process during the interviews, the most frequently discussed details related to lexical choices. Student 4, who expressed positive opinions about ChatGPT's proficiency and demonstrated strong affective engagement, described her revision process as follows:

> *I noticed that ChatGPT could effectively address lexical issues. For instance, when it comes to the word "强调", I initially translated it as "underline". However, ChatGPT corrected it to "emphasize", which I hadn't considered during my translation. I thought this word fit better with the context of news.*

In contrast, Student 1, who expressed low satisfaction with ChatGPT's feedback and considered some suggestions unnecessary, chose to retain certain original translations unchanged. The student's remarks highlight how affective engagement with ChatGPT feedback influences decision-making regarding its uptake.

The screen recordings revealed that students employed several observable revision strategies to improve the quality of their translations. Four commonly used strategies are summarised in Table 8. The most frequently used strategy was online searching, which occurred 32 times in total. Three interviewees reported using this strategy to address uncertainties about specific feedback points while simultaneously correcting self-identified issues. For instance, when Student 3 noticed that ChatGPT had revised her original phrase "pragmatic conversation" to

Table 8 Summary of revision strategies

Strategy	Occurrence
Online searching	32
Proofreading	26
Using dictionaries	18
Using corpora	3
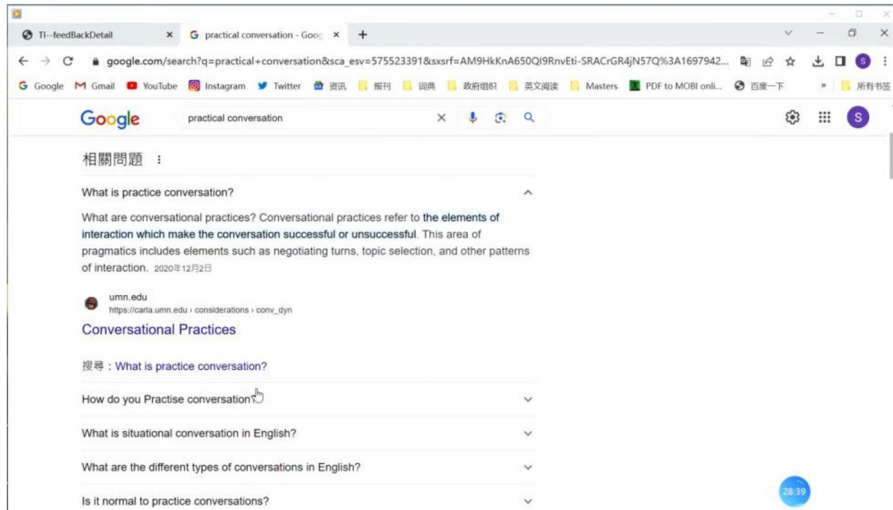
Fig. 3 Example of a search conducted by Student 3

"practical conversation," she critically assessed whether it was a commonly used collocation by searching for it on Google (Fig. 3). The second most frequently used strategy was proofreading, which occurred 26 times and was aimed at ensuring the accuracy and correctness of translations after modifications (Student 1, Student 2). According to the remarks of three interviewees, doubts about the feedback often prompted them to consult corpora and dictionaries. However, Student 2, who expressed confidence in ChatGPT's feedback, did not consult additional resources to verify the reliability of the feedback. Instead, she appeared to accept it uncritically, suggesting that her behavioural and cognitive engagement were relatively superficial.

Apart from the above-mentioned strategies, the interviews revealed that two participants would recall prior knowledge to make correction decisions and identify additional errors in the revision process. For example, Student 3 changed "extended his hope" to "in the expectation of", a change that was not suggested in ChatGPT's feedback. She elaborated on her reasoning for this correction:

When it comes to the word "希望", I initially translated it into "extend his hope". But as I was revising, I remembered something from my interpreting

classes: "extend" is usually paired with "gratitude", not "hope". So I ended up changing it to "in the expectation of".

The revision operations and strategies reveal a strong link between students' cognitive and behavioural engagement. Cognitive engagement appears to drive behavioural engagement, initiating and guiding students' actions during the revision process (Fan & Xu, 2020). This implies that students' thinking and critical evaluation directly influence their revision behaviours, emphasising the interplay between mental processes and actions in translation revision.

We categorised students' revision duration, which refers to the time spent on searching and textual modifications, into four ranges: less than 5 min, 5–20 min, 20–40 min, and more than 40 min. Figure 4 illustrates the distribution of time per category. Notably, the majority of students spent between 5–20 min (41% of participants) and 20–40 min (31% of participants) on revision. Table 9 shows that students took an average of 24:41 to complete their revisions. It is worth noting that while the minimum revision time recorded was just 01:30, which was exceptionally brief, all other students' revisions lasted longer than 08:30. The unusually short revision time of this student might be attributed to a lack of motivation or interest in the task.

Three interviewees reported spending considerable time executing search queries on the internet. Interestingly, Student 3, despite expressing satisfaction with Chat-GPT's lexical choices, spent 54:34 on revision—the longest time among all participants. Much of this effort was devoted to verifying the accuracy of ChatGPT's feedback. This reveals a discrepancy between her affective engagement and behaviours:
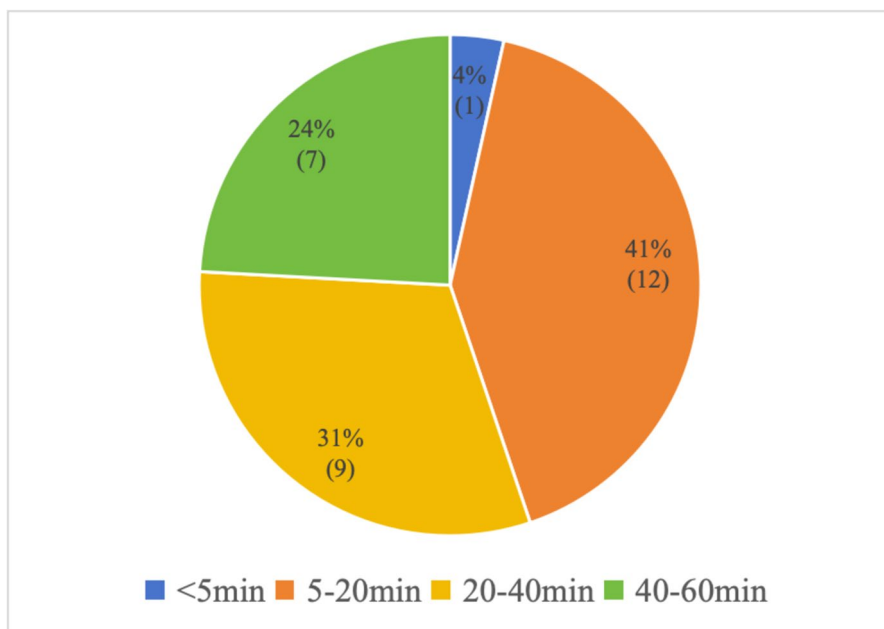

Fig. 4 Time spent on revision

Table 9 Summary of students' revision time	Mean	Min	Max	SD
	0:24:41	0:01:30	0:54:34	0:15:17

Whenever I noticed that ChatGPT's suggestions differed from my translation, I would hesitate about whether to follow them by doing some searching using Google. Besides that, I would also proofread my revised translation to catch any other potential errors.

In contrast, Student 2, who also expressed confidence in ChatGPT's feedback, accepted most of the content superficially without consulting any additional resources. This resulted in a relatively brief revision time of 09:45. The contrasting behaviours and motivations for conducting searches observed in these two students demonstrate individual differences in the mutual influence of various types of engagement.

6 Discussion

This study revealed that students generally engaged actively with ChatGPT's translation feedback across cognitive, affective, and behavioural dimensions, supporting the notion that collaboration with AI tools can positively influence language learning (Sahari et al., 2023). However, their engagement exhibited considerable complexity, reflecting intricate interactions and interdependencies both within the sub-constructs of each dimension and across the three dimensions.

Regarding cognitive engagement, participants generally found the feedback easy to understand, particularly surface-level suggestions. However, a small number of students struggled to comprehend meaning-level feedback, particularly in areas such as cohesion and coherence, and required additional guidance. This difficulty stemmed from the overly general nature of the feedback, which impeded students' ability to process it effectively. This finding aligns with Su et al.'s (2023) research, which highlighted ChatGPT's tendency to produce vague feedback. In response to the feedback, students employed various meta-cognitive strategies to manage their mental effort and plan their revision processes. These strategies demonstrated that students invested significant cognitive effort to ensure the feedback's appropriateness and the accuracy of their revised translations. Specific strategies included repeatedly reviewing feedback, proofreading their revisions, and engaging in self-correction. The substantial effort required to interpret ChatGPT's feedback is consistent with the findings of Zheng and Yu (2018), who noted that processing general feedback demanded greater linguistic competence for accurate interpretation. The mismatch between feedback comprehension and the meta-cognitive strategies employed underscores inconsistencies within the sub-constructs of cognitive engagement. Students showed varied preferences in deciding which parts to prioritise upon reviewing the feedback.

As observed in the study, most participants initially focused on checking their grades and evaluating their overall performance. This behaviour reflects a competence-driven mindset often associated with traditional classroom learning contexts (Cassidy, 2008; Elliot et al., 2018; Huguet et al., 2001; Riemer & Schrader, 2022). When addressing specific feedback content, students tended to prioritise surface-level issues. These were less cognitively demanding (Yu et al., 2019) and required lower linguistic competence compared to meaning-level issues (Chandler, 2003). Furthermore, conflicting behaviours were observed within cognitive operations, aligning with Jiang and Yu's (2022) findings, which highlighted students' reservations about relying on automated feedback. Many participants consulted additional resources to determine whether to incorporate ChatGPT's feedback into their work. This proactive engagement suggests that ChatGPT feedback encourages students to critically evaluate their doubts, thereby fostering critical thinking skills—a key competency for translators.

Students' affective engagement with ChatGPT feedback presented a complex picture, consistent with their cognitive engagement. They reported enjoying the revision process and feeling confident after reviewing the high grades and positive comments generated by ChatGPT. However, although students initially prioritised checking their grades upon receiving the feedback, some shifted their focus to the feedback content. This preference for feedback over grades may stem from the fact that ChatGPT's scoring did not compare individual translation performance with that of the entire class. This observation highlighted students' natural tendency to compare their grades with those of their peers in learning activities, reflecting findings related to cognitive engagement. Notably, students appeared to prefer critical feedback that offered more guidance for improvement, which enhanced their motivation to complete revisions. This finding contradicts the research of Ilies et al. (2007), which suggested that negative feedback could undermine students' confidence and adversely affect their mood. Overall, students held generally positive views of ChatGPT in the process of knowledge acquisition (Sallam et al., 2023). They appreciated the unique feedback structure and ChatGPT's ability to address specific issues, particularly in lexical choices and sentence structure. However, some students expressed frustration with the effectiveness of ChatGPT feedback, as not all suggestions were considered necessary. This indicates that difficulties in interpreting ChatGPT feedback in the cognitive domain can influence students' attitudes toward it.

Behaviourally, the revisions based on ChatGPT feedback were not as effective as expected. This can be attributed to previous research suggesting that learners are selective in their adoption of automated feedback, adjusting their uptake accordingly (Bai & Hu, 2017; Qian & Li, 2023; Storch & Wigglesworth, 2010). Notably, students often identified issues not detected by ChatGPT and made self-initiated changes, which may reflect their inherent scepticism about the accuracy of automated feedback, as noted by Fan and Xu (2020). This phenomenon underscores the complex relationship between cognitive engagement and behavioural engagement. Regarding specific aspects of feedback, although students reported difficulties in understanding meaning-level feedback and tended to prioritise surface-level issues, the majority of revisions were based on meaning-level feedback, which had a higher uptake rate. This disparity between behavioural and cognitive engagement aligns

with Zhang et al.'s (2023) findings. A key contributing factor may be that, while ChatGPT excels at providing contextually appropriate vocabulary, it often fails to offer sufficiently detailed and explicit suggestions on genre conventions, thereby constraining students' ability to effectively engage with surface-level feedback. This limitation supports Zheng and Yu's (2018) assertion that the feedback provider's approach can significantly shape students' revision behaviours.

Furthermore, our interview results revealed that the incorporation of feedback and decision-making processes were influenced by students' affective engagement (Yu et al., 2019). However, the analysis of revision periods and detailed revision actions suggest that the interplay among the three dimensions of engagement is complex and varies due to individual factors, such as prior knowledge, learning styles, and motivational orientations (Afifi et al., 2023). For instance, students with higher levels of prior knowledge in the subject matter may be more likely to engage cognitively with the feedback, while those with lower levels of prior knowledge may rely more heavily on affective engagement, such as their emotional responses to the feedback. A case in point is Student 3, who expressed satisfaction with the feedback but still had doubts about its reliability. Notably, she invested considerable time in comprehending the feedback, demonstrating deep cognitive engagement, which in turn led to substantial behavioural engagement. Additionally, students devoted effort to verifying the reliability of ChatGPT feedback and subsequently employed strategies to address the feedback and regulate translation accuracy. This finding aligns with Wang et al.'s (2022) study, which highlighted the impact of cognitive processing on behavioural engagement. The complex interplay among the three dimensions of engagement underscores the need for further investigation into how students interact with automated feedback, taking into account individual differences and contextual factors.

Several pedagogical implications can be drawn from this study. First, feedback itself cannot lead to learning gains unless students generate internal feedback through their own processing of the information provided (Yan & Carless, 2022). Students in this study reflected that the general nature of the feedback prevented them from processing it effectively, underscoring the importance of considering how the characteristics of feedback impact student engagement and influence their revision processes. Second, it is notable that students prefer to utilise more critical suggestions over positive feedback. Accordingly, teachers should adopt a holistic approach when designing and adjusting feedback strategies based on individual engagement and performance (Shen & Chong, 2023). Specifically, the design of feedback should ensure a balanced inclusion of both recognition of strengths and constructive critical comments. The study found that students tended to employ cognitive operations to decide whether to accept feedback and utilised meta-cognitive strategies to address feedback and make revisions. However, managing the quality of translation and processing feedback information can be challenging for students, particularly when their language proficiency and feedback literacy are limited (Zhang, 2020). To address these challenges, additional instruction should be incorporated into daily teaching, including guidance on developing linguistic skills, strategies for applying feedback effectively, and techniques for error detection (Malecka et al., 2022; Zhang, 2017). Finally, students' engagement patterns suggest that ChatGPT

performs satisfactorily on basic levels, including lexical choices and sentence structures, while showing moderate performance on higher-level issues, such as cohesion and coherence. In future translation teaching, ChatGPT feedback can be effectively integrated with teacher feedback to address basic-level concerns. This approach allows teachers to concentrate on providing more nuanced feedback, ensuring comprehensive coverage and potentially reducing their workload, especially when dealing with the practical challenges of large class sizes (Guo et al., 2024). However, to successfully integrate AI into classroom practices, it must be developed with a clear understanding of educational regulations (Kim, 2023). Educators should also enhance their expertise in prompt engineering to improve the efficacy of ChatGPT outcomes.

7 Conclusion

The findings of this study underscore the transformative potential of ChatGPT-generated feedback in the realm of translation education. By introducing AI into the feedback loop, educators can not only alleviate the heavy workload traditionally associated with manual feedback but also enhance the quality and timeliness of the feedback provided to students. This integration can create a more dynamic and responsive learning environment where students receive immediate and detailed insights into their translation work. Moreover, the study highlights the importance of fostering a robust engagement framework that encompasses cognitive, affective, and behavioural dimensions. A detailed understanding of how students interact with AI-generated feedback can inform the development of more effective pedagogical strategies. For instance, tailoring feedback to address individual learning needs and preferences can significantly enhance student satisfaction and learning outcomes. Furthermore, teachers should offer guidance to ensure students use generative AI appropriately during the learning process.

Despite these findings and the valuable insights gained, our study has some limitations that should be acknowledged. First, since the output of ChatGPT is influenced by the quality of prompts (OpenAI, 2022), the effectiveness of feedback could be improved by enhancing prompt quality. Second, we utilised ChatGPT 3.5, the most advanced version available at the time of the experiment. As technology progresses, more recent AI models should be considered to ensure the study remains current. Third, our investigation of student engagement was limited to a single experiment, which may not fully capture the complexities of engagement over time. Future studies could explore the long-term impact of ChatGPT feedback on student learning by conducting longitudinal research. Moreover, exploring the differences between ChatGPT feedback and human feedback could provide valuable insights into improving translation feedback and developing effective strategies for integrating the two sources in translation education. Further research could also investigate the potential of ChatGPT feedback across diverse learner populations, including those with varying levels of proficiency and training experience, to determine its broader applicability in translation education.

Appendix 1. Guide for the semi-structured interview

1. What actions did you take upon receiving feedback? Could you describe them in detail?
2. How did you feel after receiving your first feedback?
3. How did you feel after receiving your second feedback?
4. Please compare the feedback on the two drafts and tell me which one you think is better, and why?
5. What are your thoughts on the holistic scores and feedback?
6. Do you find the feedback helpful?
7. Is the feedback easy to understand?
8. Did you encounter any problems with the feedback?
9. When you encountered problems with the teacher's feedback, what did you do?
10. Can you identify the strengths and weaknesses of ChatGPT feedback?
11. Do you think ChatGPT feedback should be used in translation education? Why?
12. How do you think ChatGPT feedback could be utilised in translation teaching?

Appendix 2. A sample of ChatGPT-generated feedback

1. Reference translation

Looking ahead to the 2024 fiscal year, in the face of the uncertainties brought by the macro environment, we will still focus on our long-term strategic goals, steadfastly advancing globalization, strengthening our product strength, and further optimizing our store network. We will participate in global competition from two dimensions: cost leadership and product differentiation. While always adhering to the basic principle of cost-effectiveness, we will continue to produce high-quality works with IP design as a feature, making daily necessities more fashionable and trendy. At the same time, we will also actively explore the super store strategy, establishing a strong brand image in the minds of consumers through super stores, targeting big beauty and cosmetics, big toys, and big IP to create super categories, and further exploring the potential for increasing average store sales.

2. Comments and suggestions

1) The translation accurately conveys the meaning of the original text, maintaining the same information and clarity.
2) The register of the translation matches that of the source text, maintaining a formal and professional tone.
3) The composition in the target language is well done, with correct orthography and typography.
4) The vocabulary used is appropriate and rich, effectively conveying the concepts and ideas of the source text.

5) The morphosyntax is good, with proper use of verb tenses, prepositions, and other grammatical elements.
6) The cohesion of the text is well-maintained, with good use of connectors and referential elements.
7) The coherence of the text is well-established, with ideas organized and presented clearly.
8) The level of communication of the target text is high, with an overall quality that effectively conveys the intentions of the source text.
9) The appropriateness in terms of the genre's conventions and the translation's purpose and target audience is well-maintained.

3. Grade

1) Expression of the meaning of the original text: 38/40
2) Composition in the target language: 38/40
3) Level of communication of the target text: 19/20
4) Overall score: 95/100

Declarations

References

Afifi, S., Rahimi, M., & Wilson, J. (2023). Student engagement with teacher and automated written corrective feedback on L2 writing: A multiple case study. *The JALT CALL Journal, 19*(2), 216–242. https://doi.org/10.29140/jaltcall.v19n2.1041

AlGhamdi, R. (2024). Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study. *Education and Information Technologies, 29*, 18901–18926. https://doi.org/10.1007/s10639-024-12594-2

Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology, 37*(1), 67–81. https://doi.org/10.1080/01443410.2016.1223275

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education, 21*(1), 23. https://doi.org/10.1186/s41239-024-00455-4

Bruton, A. (2007). Vocabulary learning from dictionary referencing and language feedback in EFL translational writing. *Language Teaching Research, 11*(4), 413–431. https://doi.org/10.1177/1362168807080961

Cao, S., & Zhong, L. (2023). Exploring the effectiveness of ChatGPT-based feedback compared with teacher feedback and self-feedback: Evidence from Chinese to English translation. *arXiv preprint* arXiv:2309.01645. https://arxiv.org/abs/2309.01645. Accessed 16 Dec 2023.

Cassidy, S. (2008). Approaches to learning and competitive attitude in students in higher education. *Psychology of Education Review, 32*(1), 18–27. https://doi.org/10.53841/bpsper.2008.32.1.18

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*(3), 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Chen, Y. (2021). Understanding student engagement with teacher and peer feedback in a translation course for Chinese English majors [Doctoral dissertation, The Hong Kong Polytechnic University]. Hong Kong Polytechnic University.

Cheng, X., & Zhang, L. J. (2024). Examining Second Language (L2) Learners' Engagement with AWE-Teacher Integrated Feedback in a Technology-Empowered Context. *The Asia-Pacific Education Researcher, 33*, 1023–1035. https://doi.org/10.1007/s40299-024-00877-8

Cheng, X., Liu, Y., & Wang, C. (2023). Understanding student engagement with teacher and peer feedback in L2 writing. *System, 119*, 103176. https://doi.org/10.1016/j.system.2023.103176

Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y. S., Gasevic, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. In N.-S. Chen, G. Rudolph, D. G. Sampson, M. Chang, R. Kuo, & A. Tlili (Eds.), *Proceedings - 2023 IEEE International Conference on Advanced Learning Technologies, ICALT 2023* (pp. 323–325). IEEE, Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICALT58122.2023.00100

Ekin, S. (2023). Prompt Engineering For ChatGPT: A Quick Guide To Techniques, Tips, And Best Practices. *TechRxiv preprint* https://doi.org/10.36227/techrxiv.22683919.v2

Elliot, A. J., Jury, M., & Murayama, K. (2018). Trait and perceived environmental competitiveness in achievement situations. *Journal of Personality, 86*(3), 353–367. https://doi.org/10.1111/jopy.12320

Ellis, R. (2010). EPILOGUE: A Framework for Investigating Oral and Written Corrective Feedback. *Studies in Second Language Acquisition, 32*(2), 335–349. https://doi.org/10.1017/S0272263109990544

Er, E., Dimitriadis, Y., & Gašević, D. (2021). Collaborative peer feedback and learning analytics: Theory-oriented design for supporting class-wide interventions. *Assessment & Evaluation in Higher Education, 46*(2), 169–190. https://doi.org/10.1080/02602938.2020.1764490

Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing, 50*, 100775. https://doi.org/10.1016/j.jslw.2020.100775

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

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*, 8435–8463. https://doi.org/10.1007/s10639-023-12146-0

Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education, 63*, 100962. https://doi.org/10.1016/j.iheduc.2024.100962

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Investigating students' engagement in chatbot-supported classroom debates. *Interactive Learning Environments, 32*(9), 4917–4933. https://doi.org/10.1080/10494820.2023.2207181

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing, 30*, 31–44. https://doi.org/10.1016/j.jslw.2015.08.002

Han, C., & Lu, X. (2023). Can automated machine translation evaluation metrics be used to assess students' interpretation in the language learning classroom? *Computer Assisted Language Learning, 36*(5–6), 1064–1087. https://doi.org/10.1080/09588221.2021.1968915

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': Investigating the concept of student engagement with feedback. *Oxford Review of Education, 37*(4), 543–560. https://doi.org/10.1080/03054985.2011.604951

Hattie, J. (1992). Measuring the Effects of Schooling. *Australian Journal of Education, 36*(1), 5–13. https://doi.org/10.1177/000494419203600102

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports, 13*(1), 18617. https://doi.org/10.1038/s41598-023-45644-9

Huguet, P., Dumas, F., Monteil, J. M., & Genestoux, N. (2001). Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology, 31*(5), 557–578. https://doi.org/10.1002/ejsp.81

Hurtado Albir, A., & Taylor, P. (2015). The acquisition of translation. competence competences, tasks, and assessment in translator training. *Meta, 60*(2), 256–280. https://doi.org/10.7202/1032857ar

Ilies, R., De Pater, I. E., & Judge, T. (2007). Differential affective reactions to negative and positive feedback, and the role of self-esteem. *Journal of Managerial Psychology, 22*(6), 590–609. https://doi.org/10.1108/02683940710778459

Jiang, L., & Yu, S. (2022). Appropriating automated feedback in L2 writing: Experiences of Chinese EFL student writers. *Computer Assisted Language Learning, 35*(7), 1329–1353. https://doi.org/10.1080/09588221.2020.1799824

Jørgensen, B. M. (2018). Investigating non-engagement with feedback in higher education as a social practice. *Assessment & Evaluation in Higher Education, 44*(4), 623–635. https://doi.org/10.1080/02602938.2018.1525691

Kim, J. (2023). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies, 29*, 8693–8724. https://doi.org/10.1007/s10639-023-12109-5

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing, 44*, 100450. https://doi.org/10.1016/j.asw.2020.100450

Lee, S.-M., Yang, Z., & Wu, J. G. (2024). Live, play, and learn: Language learner engagement in the immersive VR environment. *Education and Information Technologies, 29*, 10529–10550. https://doi.org/10.1007/s10639-023-12215-4

Lee, T. K. (2023). Artificial intelligence and posthumanist translation: ChatGPT versus the translator. *Applied Linguistics Review*, *15*(6), 2351–2372. https://doi.org/10.1515/applirev-2023-0122

Li, X., & Ke, P. (2022). An empirical study of peer feedback in translation didactics: Quality, response and student perceptions. *Assessment & Evaluation in Higher Education, 47*(8), 1231–1244. https://doi.org/10.1080/02602938.2022.2044013

Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education, 27*(7), 908–922. https://doi.org/10.1080/13562517.2020.1754784

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

OpenAI. (2022). ChatGPT: Optimizing language models for dialogue. *OpenAI.* https://chatgpt.r4wand.eu.org/. Accessed 10 Nov 2023.

Pan, J., Chen, H., & Yuan, S. (2023). A comparative study of the engagement with written corrective feedback of Chinese private college students. *Asian-Pacific Journal of Second and Foreign Language Education, 8*(1), 18. https://doi.org/10.1186/s40862-023-00191-8

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics, 36*, 50–72. https://doi.org/10.1017/S0267190515000094

Qian, J., & Li, D. (2023). Toward a better understanding of student engagement with peer feedback: A longitudinal study. *International Review of Applied Linguistics in Language Teaching.* https://doi.org/10.1515/iral-2023-0108

Riemer, V., & Schrader, C. (2022). Playing to learn or to win? The role of students' competition preference on self-monitoring and learning outcome when learning with a serious game. *Interactive Learning Environments, 30*(10), 1838–1850. https://doi.org/10.1080/10494820.2020.1752741

Sahari, Y., Al-Kadi, A. M. T., & Ali, J. K. M. (2023). A cross sectional study of ChatGPT in Translation: Magnitude of use, attitudes, and uncertainties. *Journal of Psycholinguistic Research, 52*(6), 2937–2954. https://doi.org/10.1007/s10936-023-10031-y

Sallam, M., Salim, N. A., Barakat, M., Al-Mahzoum, K., Al-Tammemi, A. B., Malaeb, D., Hallit, R., & Hallit, S. (2023). Assessing health students' attitudes and usage of ChatGPT in Jordan: Validation study. *JMIR Medical Education, 9*, e48254. https://doi.org/10.2196/48254

Shen, R., & Chong, S. W. (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. https://doi.org/10.1080/02602938.2022.2072468

Shen, W.-W., Tsai, M.-H. M., & Lin, J.-M. (2017). The ideal and reality of implementing technology into English language teaching: A case study of using a peer review system. In *International Symposium on Emerging Technologies for Education* (Vol. 10676, pp. 391–399). Springer International Publishing. https://doi.org/10.1007/978-3-319-71084-6_44

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*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case Studies. *Studies in Second Language Acquisition, 32*(2), 303–334. https://doi.org/10.1017/S0272263109990532

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System, 91*, 102247. https://doi.org/10.1016/j.system.2020.102247

Wang, B., Yu, S., Zheng, Y., & Teo, T. (2022). Student engagement with teacher oral feedback in EFL university classrooms. *Language Teaching Research*, 136216882211057. https://doi.org/10.1177/13621688221105772

Washbourne, K. (2014). Beyond error marking: Written corrective feedback for a dialogic pedagogy in translator training. *The Interpreter and Translator Trainer, 8*(2), 240–256. https://doi.org/10.1080/1750399X.2014.908554

Winstone, N. E., Nash, R. A., Rowntree, J., & Parker, M. (2017). 'It'd be useful, but I wouldn't use it': Barriers to university students' feedback seeking and recipience. *Studies in Higher Education, 42*(11), 2026–2041. https://doi.org/10.1080/03075079.2015.1130032

Yan, Z., & Carless, D. (2022). Self-assessment is about more than self: The enabling role of feedback literacy. *Assessment & Evaluation in Higher Education, 47*(7), 1116–1128. https://doi.org/10.1080/02602938.2021.2001431

Yu, S., Zhang, Y., Zheng, Y., Yuan, K., & Zhang, L. (2019). Understanding student engagement with peer feedback on master's theses: A Macau study. *Assessment & Evaluation in Higher Education, 44*(1), 50–65. https://doi.org/10.1080/02602938.2018.1467879

Yu, S., Zhang, Y., Zheng, Y., & Lin, Z. (2020). Written corrective feedback strategies in English-Chinese translation classrooms. *The Asia-Pacific Education Researcher, 29*(2), 101–111. https://doi.org/10.1007/s40299-019-00456-2

Zhang, Z. (2017). Student engagement with computer-generated feedback: A case study. *ELT Journal, 71*(3), 317–328. https://doi.org/10.1093/elt/ccw089

Zhang, R., Zou, D., & Cheng, G. (2023). Learner engagement in digital game-based vocabulary learning and its effects on EFL vocabulary development. *System, 119*, 103173. https://doi.org/10.1016/j.system.2023.103173

Zhang, Z. (Victor), & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, *36*, 90–102.https://doi.org/10.1016/j.asw.2018.02.004

Zhang, Z. (Victor). (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing, 43*, 100439. https://doi.org/10.1016/j.asw.2019.100439

Zheng, Y., & Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing, 37*, 13–24. https://doi.org/10.1016/j.asw.2018.03.001

Zheng, Y., Zhong, Q., Yu, S., & Li, X. (2020). Examining Students' Responses to Teacher Translation Feedback: Insights From the Perspective of Student Engagement. *SAGE Open, 10*(2), 215824402093253. https://doi.org/10.1177/2158244020932536

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.