Learning and Motivation 91 (2025) 102158

Contents lists available at ScienceDirect

Learning and Motivation

journal homepage: www.elsevier.com/locate/l&m

Enhancing EFL writing with visualised GenAI feedback: A cognitive affective theory of learning perspective on revision quality, emotional response, and human-computer interaction

Bin Zou [a,1], Chenghao Wang [a,b,*,2], Huimin He [c,3], Congxin Li [a], Erick Purwanto [d], Ping Wang [e]

[a] Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, China
[b] Department of English, School of Arts, Univeristy of Liverpool, UK
[c] School of Languages, Xi'an Jiaotong-Liverpool University, China
[d] Computing Department of School of Advanced Technology, Xi'an Jiaotong-Liverpool University, China
[e] Institute of Leadership and Education Advanced Development, Xi'an Jiaotong-Liverpool University, China

ARTICLE INFO

Keywords:
Generative AI visualised feedback
EFL writing
Revision quality
Human-computer interaction
Emotional response
Willingness to write
CATLM

ABSTRACT

With the rapid development of large language models and natural language processing technologies, Generative AI (GenAI) chatbots have offered new opportunities for supporting EFL learners in writing and revision. However, existing GenAI chatbots' output is predominantly presented in plain, text-only formats, which may impose a high cognitive load and trigger negative emotional responses. Despite growing interest in GenAI-assisted writing instruction, limited research has examined how visual enhancements to GenAI chatbot output might improve revision outcomes and learners' emotional experiences. Grounded in the Cognitive-Affective Theory of Learning with Media, this study employed a self-developed GenAI-powered writing chatbot to investigate the effects of visualised feedback on EFL learners' writing performance and emotional responses during revision. A 2 (time: pre-test and post-test) × 2 (feedback mode: visualised vs. non-visualised) quasi-experimental design was employed. Group A ($N = 30$) received standard text-only feedback, while Group B ($N = 30$) received visualised feedback incorporating colour variation, tabular formatting, and bolded text. Results from pre-and post-tests and questionnaires indicated that visualised feedback significantly improved coherence and cohesion in learners' writing, reduced negative emotions, and resulted in lower cognitive load. These findings offer practical implications for language educators, learners, and developers, highlighting the critical role of how AI-generated content is presented in building emotionally supportive and cognitively effective GenAI-assisted learning environments.

* Correspondence to: School of Humanities and Social Sciences (HSS) Building, Xi'an Jiaotong-Liverpool University, Wuzhong District, Suzhou 215000, China.

E-mail addresses: Bin.Zou@xjtlu.edu.cn (B. Zou), dancerluo@outlook.com (C. Wang), Huimin.He@xjtlu.edu.cn (H. He), Congxin.Li@xjtlu.edu.cn (C. Li), Erick.Purwanto@xjtlu.edu.cn (E. Purwanto), Ping.Wang@xjtlu.edu.cn (P. Wang).

[1] ORCID: https://orcid.org/0000-0002-4863-0998
[2] ORCID: https://orcid.org/0009-0009-5655-3740
[3] ORCID: https://orcid.org/0009-0006-7261-5284

https://doi.org/10.1016/j.lmot.2025.102158
Received 8 May 2025; Received in revised form 22 June 2025; Accepted 23 June 2025
Available online 27 June 2025
0023-9690/

1. Introduction

In recent years, the emergence of Generative Artificial Intelligence (GenAI) has transformed the landscape of foreign language education (Du & Alm, 2024; Feng et al., 2025; Huang et al., 2024; Wang, Wang & Zou, 2024; Wang & Xue, 2024). Its powerful text generation and analysis capabilities have enhanced English as a Foreign Language (EFL) writing support (Guo & Wang, 2023; Hwang et al., 2023) while simultaneously contributing to more dynamic and engaging AI–human interaction experiences (Du & Daniel, 2024; Liu et al., 2024). With advanced natural language processing capabilities, GenAI-powered chatbots such as ChatGPT have demonstrated considerable potential in assisting EFL learners with automated feedback, error correction, and content or image generation while also offering effective support in the learning process (Feng et al., 2025; Wang, Wang, Zou, 2024; Zou et al., 2024). Unlike peer or teacher feedback, GenAI chatbots provide 24/7 accessibility and low-pressure interactions while also avoiding the overly positive, face-saving tendencies often found in peer feedback (Cai et al., 2023; Yan, 2023; Zou et al., 2024). Additionally, using GenAI to provide writing feedback can alleviate the substantial workload on teachers and expedite the process of providing feedback (Link et al., 2022; Tran, 2025) so that students can receive timely suggestions on error corrections (Alsofyani & Barzanji, 2025). Moreover, GenAI chatbots are inclined to deliver global-level feedback through summaries, an element that was considered helpful for providing a positive feedback experience but was deficient in the teacher's feedback (Guo & Wang, 2023; Lu et al., 2024; Wei & Li, 2023).

In light of these favourable features, GenAI chatbot feedback has been widely shown to enhance EFL learners' writing skills and overall writing quality (Abdullayeva & Musayeva, 2023; Huang & Tan, 2023; Hwang et al., 2023; Mahapatra, 2024). For instance, Huang and Tan (2023) found that the GenAI chatbot (ChatGPT) effectively helped learners improve text coherence and cohesion through instant textual analysis and sentence suggestions. GenAI chatbots, by providing efficient and targeted comments on content, language, and structure, serve as a valuable complement to teacher feedback (Boudouaia et al., 2024; Huang & Hu, & Jin, 2024; Luo et al., 2025; Su et al., 2023). However, as GenAI output relies on user prompts (input), EFL students often experience high cognitive load during prompt design, potentially hindering their learning experience (Woo et al., 2024). Moreover, excessive and verbose outputs, combined with the inherent cognitive demands of writing and revision, also contributed to increased cognitive load (Flower & Hayes, 1981; Zou et al., 2024).

On the basis of the Cognitive Load Theory (CLT) (Sweller, 2008), Mayer (2005) established the Cognitive Theory of Multimedia Learning (CTML), examining how people process and understand information delivered through various media formats. According to CTML, active cognitive engagement is the key to effective multimedia learning, so instructional designers should try to reduce the extraneous cognitive load on learners in multimedia environments, thereby enhancing their academic achievement (Mayer, 2005; Woo et al., 2024). This implies that optimising interactions and output as feedback in GenAI-assisted writing and revision can help reduce cognitive overload and, in turn, enhance writing quality. Furthermore, to extend the CTML, Moreno (2006) proposed Cognitive-Affective Theory of Learning with Media (CATLM, see Fig. 1). The model illustrates how sensory input from various instructional media is processed through working memory and ultimately stored in long-term memory, more importantly, highlighting the mediating role of affect (emotion). Rather than treating emotion as peripheral, CATLM (Moreno, 2006) positions it as a core factor that interacts with cognitive processing to influence attention, information selection, integration, and ultimately learning outcomes. In other words, the CATLM extends the CTML by incorporating emotional factors, asserting that a combination of emotional, motivational, and metacognitive elements influences how learners process content in multisensory learning environments. As noted by Gottfried (1990) and Moreno et al. (2001), emotional factors can help regulate the learning process by increasing or decreasing learners' cognitive engagement, as shown in Fig. 1. As a result, adjusting how GenAI chatbot outputs are presented, the format of feedback, to reduce cognitive load could influence learners' emotion, which in turn may have a positive impact on their academic performance. Grounded in the CATLM, research on emotional design focuses on using design elements, such as verbal and visual

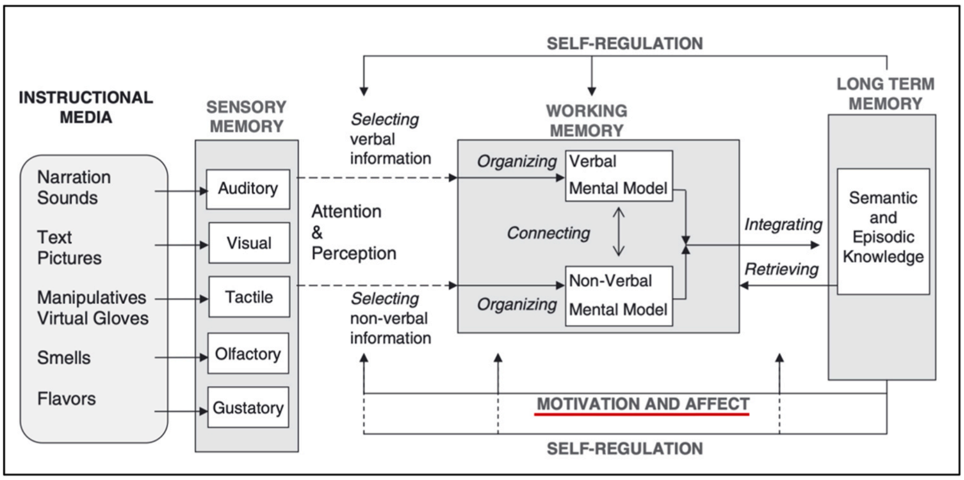

Fig. 1. Cognitive-Affective Theory of Learning with Media (Moreno, 2006, p. 151).

factors, to affect learners' emotions and promote learning by changing the way information is presented and interacted in multimedia learning environments (Li et al., 2025; Plass & Kaplan, 2016; Tan et al., 2021; Wang & Zou, 2025; Wang & Li, 2025). Inspired by the CATLM, researchers have begun to explore how visual features, including colour, structure, clarity, and layout, can shape learners' learning experiences, with emotionally supportive and aesthetically engaging designs shown to promote learner well-being and foster sustained engagement in the learning process (Li et al., 2025; Nie et al., 2024; Plass & Kaplan, 2016; Tan et al., 2021; Wang & Li, 2025). However, in the GenAI-assisted language learning contexts, limited research has explored how the emotional design of GenAI chatbot feedback influences learning achievement.

According to the CATLM (Moreno, 2006), positive emotional design is based on the premise that learners' emotional states significantly influence cognitive processing and learning outcomes. Positive emotional design elements such as warm or bright colours and human-like features have been shown to be effective in promoting both emotional engagement and cognitive processing (Mayer & Estrella, 2014; Um et al., 2011). Previous studies verified that such design could contribute to learners' positive learning experience, comprehension and/or transfer performance (Brom et al., 2018; Plass et al., 2014; Um et al., 2011). For example, Plass et al. (2014) revealed that warm colours and round face-like shapes induced positive emotions and facilitated comprehension. Similarly, Li et al. (2025) found that visualised diagrams can assist students in understanding abstract concepts, thereby deepening their knowledge comprehension, enhancing cognitive outcomes, and increasing learning motivation and engagement. Park et al. (2015) also found that positive emotional states and emotional design together led to improved learning outcomes and longer attention to relevant text content. The effectiveness of positive emotional design on learning outcomes may also depend on individual differences, such as culture, prior knowledge, age, and academic discipline (Münchow & Bannert, 2019; Stárková et al., 2019; Wong & Adesope, 2021). In the GenAI-assisted EFL writing context, positive emotional design of feedback warrants further investigation, as learners' emotional responses may interact with linguistic proficiency, ultimately shaping their engagement and learning outcomes (Li et al., 2025; Wang & Li, 2025; Wang, Li, & Zou, 2025; Wang et al., 2024; Wang & Zou, 2025; Wang & Xue, 2024). Moreover, most GenAI chatbots currently available on the market offer limited visual variation in information presentation, which may increase cognitive load and limit their capacity to elicit positive emotional responses.

To date, many EFL researchers have examined various learning emotions and their impact on engagement and achievement (Wang, 2023, 2023; Wang et al., 2025; Wang et al., 2025; Derakhshan et al., 2025), further echoing the CTML and CATLM, which emphasise the crucial role of emotions in shaping learning experiences and memory retention. For example, foreign language anxiety (FLA) occurs when learners are frustrated due to their inability to express themselves and to connect genuinely with others because of the constraints imposed by the new language (Horwitz, 2017), thereby generating task-irrelevant thoughts and diminishing the cognitive resources available for completing tasks (Pekrun & Perry, 2014). As a result, higher FLA leads to a lower willingness to communicate in a foreign language (MacIntyre, 2017; Wang et al., 2021) and weakens EFL achievement (Botes et al., 2020; Tanielian, 2017). Another negative emotion, boredom, is commonly described as a psychological condition when individuals perceive a lack of challenge or feel disengaged from a task not holding their interest (Macklem, 2015). It is usually accompanied by loss of meaning, inattention and decreased engagement (Wang, 2023, 2023; Wang et al., 2025), and empirical studies have verified the negative effects of boredom on EFL learners' academic performance and learning motivation (Li & Han, 2022; Wang, 2023, 2023). On the other hand, there has been growing interest in the role of positive emotions in EFL (Prior, 2019). Enjoyment is considered as one of the positive emotions that EFL learners most frequently experience (Wang & Wang, 2024; Wang et al., 2025). Dewaele and MacIntyre (2016) defined Foreign Language Enjoyment (FLE) as the positive feelings experienced by learners after they successfully navigate learning challenges, finish academic assignments, and fulfil their psychological needs during the process of acquiring a foreign language. It can boost resource development in language learning because of its ability to expand individuals' viewpoints and help them take in more during the language acquisition process (MacIntyre & Gregersen, 2012a, 2012b), leading to an increase in engagement and academic achievement in EFL study (Dong et al., 2022; Shao et al., 2020; Wang, 2023). Additionally, self-confidence plays a crucial role in a learner's readiness to engage in communication in a foreign language (MacIntyre et al., 1998), and having a certain level of self-confidence is essential for developing the ability for foreign language communication (Brown, 1994). In addition to FLE and self-confidence, willingness to write is another positive emotion which significantly affects EFL learners' writing performance, and proficient writers who possess a sufficient willingness to write are able to convey meaning in the written text more accurately and effectively (Rafiee & Abbasian-Naghneh, 2020; Sharrad & Faravani, 2021). These findings also support the emphasis of CATLM (Moreno, 2006) on the interplay between cognition and emotion in effective learning. With the rise of GenAI chatbots, increasing attention has been paid to their emotional impact on EFL learners (Cong-Lem et al., 2025).

Studies have shown that GenAI-supported instruction can reduce FLA (Ghafouri, 2024; Hawanti & Zubaydulloevna, 2023; Wang & Wang, Pan, et al., 2024) and enhance L2 motivation (Huang & Mizumoto, 2025; Wei, 2023; Xiao & Zhi, 2023) through personalised support and positive feedback. Research on EFL writing further suggests that GenAI chatbots alleviate writing anxiety by offering scaffolds in the writing process (Barrot, 2023; Yan, 2023; Zhao, 2023). Specifically, Teng (2024) found that ChatGPT's feedback enhanced learners' self-efficacy and engagement, though it focused on perceptions rather than specific emotions. Mohammed and Khalid (2025) addressed this by using psycho-emotional scales, reporting improvements in peace of mind, emotional intelligence, and writing proficiency. Beyond emotions and cognitive load, learners' satisfaction with the technology and their perception of its human-likeness (e.g., teacher-like qualities) also influence their acceptance of and willingness to continue using the tool, both of which are closely linked to sustained engagement with educational technologies (Tram et al., 2024; Huang & Zou, 2024). In cognitively demanding writing revision tasks, it is particularly important to address these dimensions and explore how positive emotional design can enhance the overall HCI experience. A well-designed user experience is often accompanied by positive emotional responses and reduced cognitive burden, thereby encouraging deeper and more sustained engagement in technology-enhanced learning. However, how different GenAI feedback designs affect learner emotions remains underexplored, particularly from a Human-Computer

Interaction (HCI) perspective that considers how visual design and interactivity shape emotional engagement. In cognitively demanding writing revision tasks (Woo et al., 2024), it is especially important to address and enhance key dimensions that shape learners' HCI experience. A high-quality user experience is often accompanied by positive emotional responses and reduced cognitive burden, which in turn fosters deeper and more sustained engagement in technology-enhanced learning (Huang & Zou, 2024; Wang & Wang, Pan, et al., 2024). Therefore, exploring the impact of visualised GenAI writing feedback on learners' emotional and interactional experiences, as well as their academic performance, is a worthwhile direction for further investigation, particularly from the perspective of the CATLM (Moreno, 2006). This framework underscores the role of positive emotional design in activating sensory pathways (e.g., visual stimuli), which influence emotion and affect, thereby facilitating the processing and internalisation of knowledge.

Hence, this study investigates how visualised GenAI chatbot feedback, featuring colour-coded highlights, textual emphasis, and structured formatting, influence EFL learners' revision quality, emotional and interactional responses, including willingness to write, enjoyment, confidence, anxiety, boredom, cognitive load, perceived teacher-likeness, and user satisfaction. Two chatbot versions were developed based on Moreno's (2006) CATLM to deliver teacher-like feedback in either a visualised or non-visualised format using the same interface. A quasi-experimental design with a 2 (feedback type: visualised vs. non-visualised) × 2 (time: pre-test vs. post-test) structure was employed to address the following two research questions:

1. Based on the CATLM, how does visualised GenAI text feedback influence EFL learners' revision quality?
2. Grounded in the CATLM, how does visualised GenAI text feedback influence EFL learners' willingness to write, emotions and human-computer interaction experience?

2. Method

2.1. Participants and context

The participants in this study were Chinese EFL learners enrolled in an English for Academic Purposes (EAP) course designed for first-year undergraduate students at a Sino-British university in China. These students were placed in the intermediate pathway of the course based on their placement test results upon entering the university (CEFR levels B1 + to B2-). One assessment component of the course was a 300-word timed argumentative essay, modelled on IELTS Academic Writing Task 2, which formed part of the final exam at the end of the student's first academic year. At the time of data collection, at the end of the first semester, the participants had received instruction on the structure and techniques of argumentative essays and had completed a formative writing assessment, where they engaged in a timed writing task in class and received feedback from their lecturers. In addition to the EAP course assessment, most students were also preparing for the IELTS exam, either to strengthen their applications for future Master's

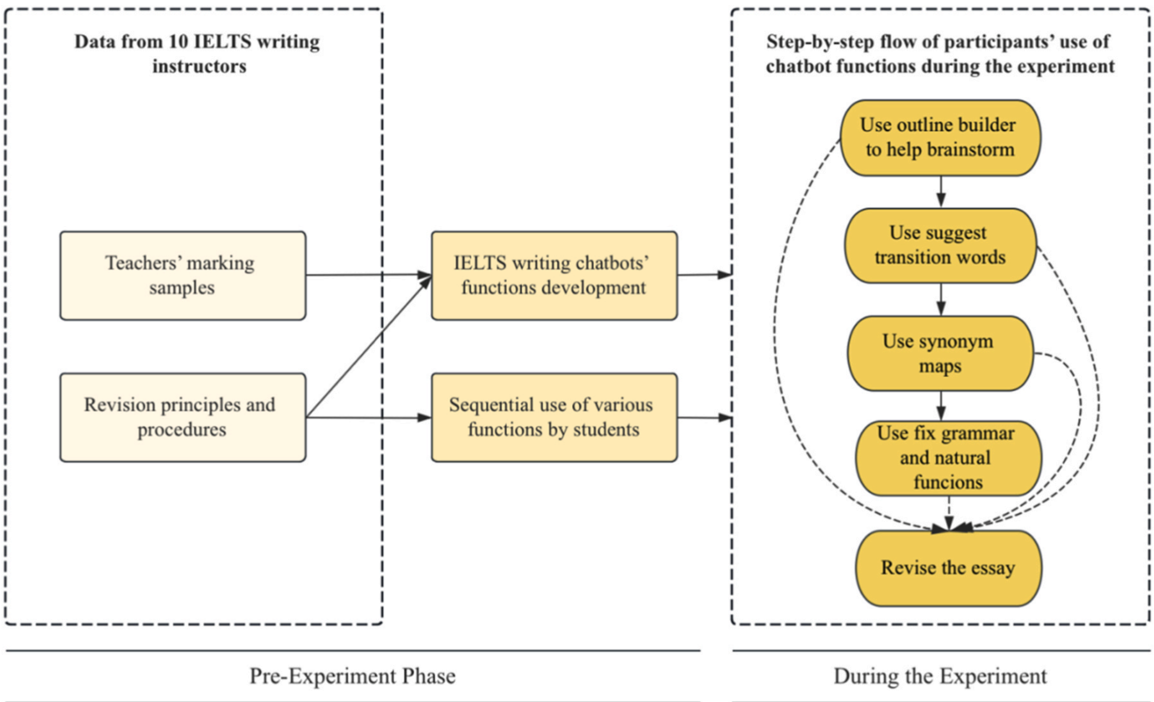

Fig. 2. Development and implementation of the Chatbot's feedback flow.

programmes or to obtain formal language certification. Overall, the participants demonstrated a generally positive and optimistic attitude toward taking part in the experiment.

2.2. Instruments

2.2.1. IELTS Writing Chatbots with different feedback formats

To develop a chatbot capable of emulating IELTS writing instructors in providing feedback on Task 2 essays and supporting learners through the revision process, we first invited ten experienced IELTS writing teachers to share samples of student essays with their corresponding marking, evaluative comments, and detailed revision principles and procedures. Based on this data, we identified a representative sample flow of teacher-led feedback and revision guidance, which served as the foundation for the chatbot's function design and experiment design (see Figs. 2 and 6). This can bridge the teacher-informed feedback logic and AI-mediated revision practices.

The teachers' revision process largely followed the four assessment criteria of the IELTS writing rubric (British Council, 2023). They typically began by reviewing the main body paragraphs to determine whether the arguments were off-topic or partially misaligned with the task requirements *(Task Response)*. They then proceeded from the beginning of the essay, identifying grammatical errors, sentence complexity, word choice, spelling mistakes, overly colloquial expressions, and unnecessary wordiness *(Lexical Resource and Grammatical Range and Accuracy)*. At the same time, they assessed coherence and conciseness, checking for missing or misused cohesive devices at both the sentence and paragraph levels *(Coherence and Cohesion)*.

Subsequently, based on the official IELTS assessment criteria and the above revision process, we utilised a large language model and hypertext markup language to develop the IELTS writing chatbots that support automated scoring and provide various revision suggestions. The interface is structured with the writing section on the left and interactive feedback on the right (see Fig. 3 & 4). The key distinction lies in the feedback presentation: the non-visualised Group A received feedback solely in plain text (see Fig. 3), similar to output generated by chatbots such as ChatGPT, Gemini, Kimi, and Deepseek, whereas visualised Group B received feedback with colour changes (see Fig. 4), strikethroughs, mind maps, bold text, and tables enhancing clarity. A comprehensive comparison of all functions is provided in Appendix 1. The visualised version is closely aligned with the CATLM, incorporating principles of positive emotional design into the chatbot. Its aim is not only to guide learners through the revision process but also to elicit positive emotional responses through clear visual cues and structured suggestions, thereby enhancing academic outcome in a GenAI-assisted learning environment.

The main feedback function includes score feedback (Grade), writing outlines (Outline Builder), transition word suggestions, synonym replacements (Synonym Maps), grammar corrections (Fix Grammar), and fluency and coherence enhancements (Natural) (see Fig. 3). These functions closely mirror the steps in the revision practices of human teachers (interviewees), as explained in Fig. 2. Specifically, for the visualised one, the outline is represented in a tree-like structure; English words in red highlight errors (e.g., incorrect words), while blue and green indicate suggested corrections or improvements (see Appendix 1 for more examples).

2.2.2. Revision quality measurement: pre-score and post-score

The primary objective of this study was to examine the impact of visual feedback on the effectiveness of writing revisions.

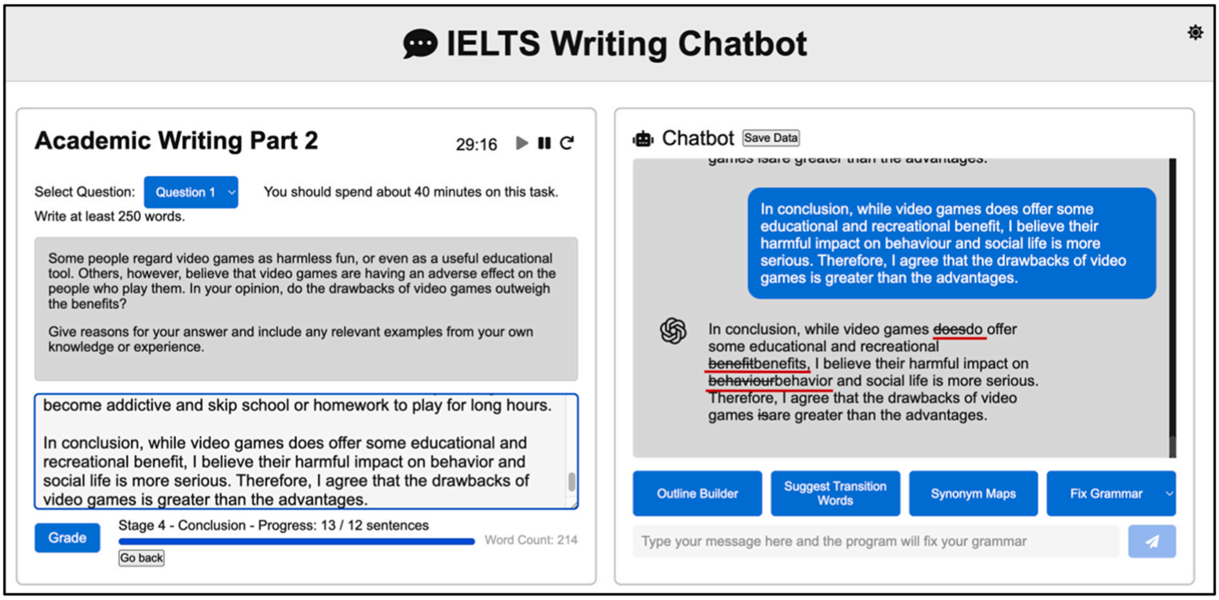

Fig. 3. Group A: Chatbot without visualised feedback (using the "Fix Grammar" function as an example).

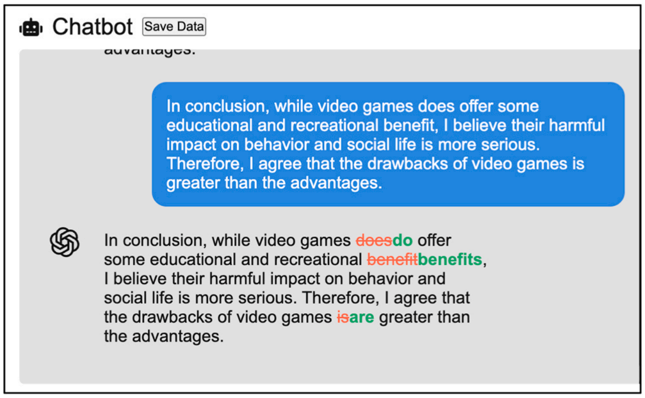

Fig. 4. Group B: Chatbot with visualised feedback (using the "Fix Grammar" function as an example).

Participants first completed an IELTS Writing Task 2 essay of approximately 300 words and then clicked the "Grade" button in the chatbot interface (see Fig. 4). The "Grade" function was developed and trained on official IELTS writing samples representing each band level, enabling it to produce band-aligned assessments and targeted feedback. The score generated by the chatbot at this point was recorded as the pre-score. Subsequently, participants revised their essays with step-by-step guidance from the chatbot (see Fig. 6), and the chatbot-generated score for the revised version served as the post-score. The chatbot's "Save Data" function automatically recorded all student scores and essays (see Fig. 5).

To evaluate the reliability of GenAI chatbot-generated grades, two experienced IELTS writing instructors independently rated 60 writing samples based on the four official IELTS writing scoring dimensions. Inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC[2,1]), a widely used measure of absolute agreement for continuous data (Koo & Li, 2016), and all dimensions demonstrated excellent consistency (ICC = 0.809–0.896). The average of the two human ratings was then compared with the AI-generated scores using the same method, revealing similarly high consistency (ICC = 0.858–0.961). These results indicate a strong level of agreement between the AI and human raters.

2.2.3. Pre- and post-survey

The pre-survey consisted of two main sections. The first section collected participants' background information, including gender, age, major, and English writing proficiency. The second section used a five-point Likert scale (1 = Strongly Disagree / Very Unwilling, 5 = Strongly Agree / Very Willing) to assess willingness to write (WTW) (5 items, adapted from Rafiee & Abbasian-Naghneh, 2020) and foreign language writing emotions, including ENY (5 items), ANX (8 items), boredom (4 items), and CON (4 items), adapted from MacIntyre et al. (1998), Wang and Wang (2024) and Wang et al. (2025).

The post-survey's emotions section mirrored that of the pre-survey, with the context incorporating feedback support provided by the chatbot. Additionally, the post-survey assessed participants' cognitive load (CL) (5 items), teacher likeness (TL) (5 items) and satisfaction (SA), adapted from Li and Wang (2024), Tram et al. (2024) and Huang and Zou (2024), respectively.

2.3. Data collection

Data were collected through a 2 (pre-test/post-test) × 2 (visualised/non-visualised) experimental design, with participants

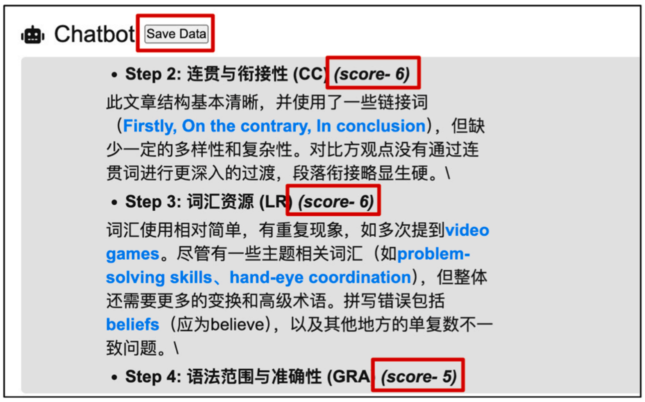

Fig. 5. The IELTS writing score generated by the chatbot's Grade function (Group B).

randomly and equally assigned to the visualised and non-visualised feedback groups. The experiment was conducted in two computer labs; the procedure followed the flow shown in Fig. 6. Specifically, after entering the lab equipped with different chatbots, participants first signed an informed consent form, followed by the completion of a pre-survey. They then watched a training video that introduced the various functions of the chatbot, including usage steps and important considerations. Participants were encouraged to ask any questions regarding unclear or unfamiliar aspects. Once all questions were addressed, they proceeded to write an argumentative essay on the computer and click "Grade" for the initial assessment. Following the sequence outlined in the right part of Fig. 6, participants used each function step by step to revise their essays. After making the revisions, they clicked "Grade" again for the post-test assessment. Finally, participants completed the post-survey.

2.4. Data analysis

We first conducted a descriptive analysis using SPSS. The skewness and kurtosis values of each were both within the ranges of |3| and |8|. Additionally, the Q-Q and P-P plots further showed that the data points were generally close to the diagonal line, indicating a normal distribution. Subsequently, for the analysis of pre-and post-scores, WTW, and foreign language writing emotions, we employed a mixed-ANOVA to examine the effects of time (within-subject effect), group (between-subject effect), and the interaction between time and group using RStudio. This analysis systematically reveals the differential impacts of various feedback modes on revision quality and writing experience. If the interaction effect was not significant, the post hoc test with the Bonferroni correction and pairwise ~ Group | Time was applied to compare two groups at each time point. This approach avoided unnecessary pairwise comparisons, enhancing the clarity and statistical efficiency of the analysis. However, if the interaction effect was significant, pairwise ~ Group * Time was adopted to examine the interaction effect fully. For CL, TL, and SA, which were measured only at the post-test, the comparison between the two groups was conducted using an independent samples t-test since it is appropriate for single-time-point comparisons, avoids unnecessary complexity, and ensures a valid statistical approach without violating assumptions.

3. Results

3.1. Descriptive results

Firstly, the descriptive analysis indicates that students in both groups who revised their IELTS Writing Task 2 essays using GenAI feedback showed improvements in willingness to write, enjoyment, and confidence, all of which exceeded 4 in the post-test (see Table 1 and Fig. 7), reflecting a more positive attitude toward writing tasks and increased confidence in revising their essays. Meanwhile, the reductions in anxiety and boredom suggest that the feedback process helped alleviate writing-related stress. Secondly,

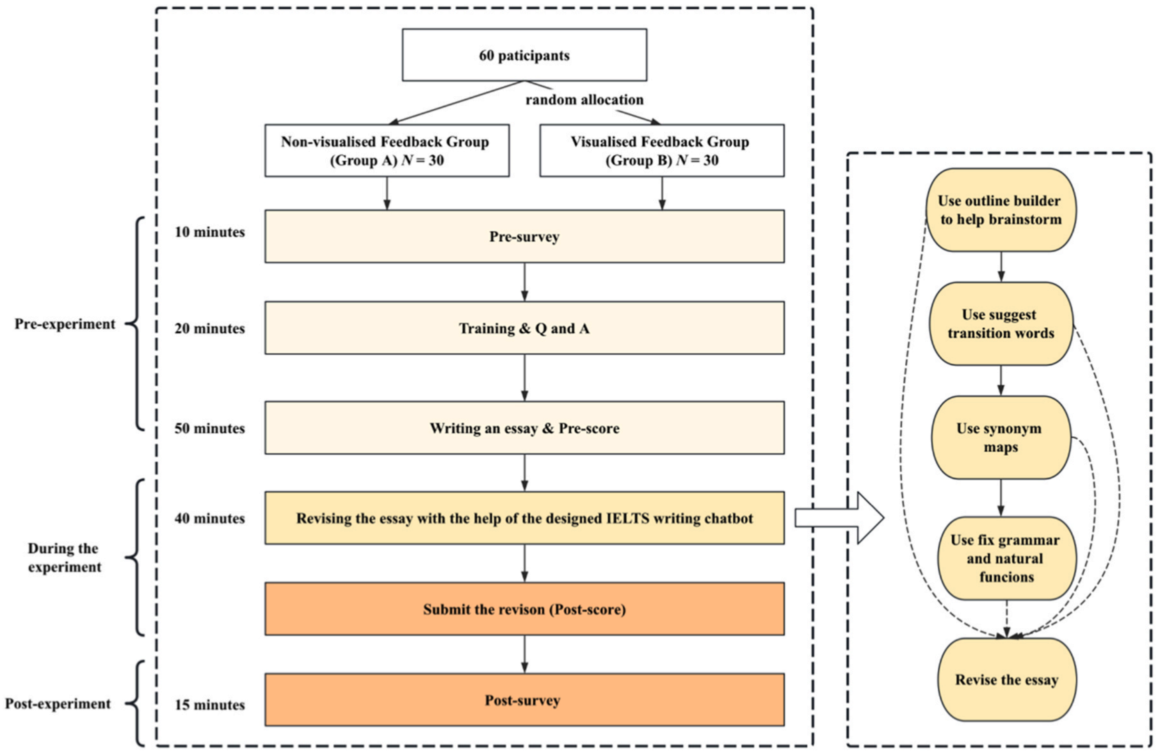

Fig. 6. Experiment flowchart and revision procedures.

in terms of human-computer interaction experience, Group A reported a higher CL ($M = 3.173$, $SD = 0.823$) than Group B ($M = 2.527$, $SD = 0.933$), suggesting that visualised feedback resulted in a lower cognitive load during the GenAI-assisted writing revision process. Regarding TL and SA, both groups showed similar and positive perceptions of teacher likeness (Group A: $M = 3.980$, $SD = 0.880$; Group B: $M = 3.973$, $SD = 0.822$) and satisfaction (Group A: $M = 4.347$, $SD = 0.575$; Group B: $M = 4.233$, $SD = 0.561$), indicating that both feedback modes effectively simulated the teacher feedback process and resulted in a high level of satisfaction with the feedback. Thirdly, in terms of revision quality, Group B demonstrated significant improvements across all five IELTS writing criteria, with a generally greater increase in scores compared to Group A. Notably, in the area of CC, Group B's score rose significantly from 5.9 to 7.233 after revision, whereas Group A's score increased from 5.8 to 6.633 (see Fig. 8). This discrepancy highlights the more pronounced effect of visualisation in enhancing the logical structure and cohesion of the writing. The specific size of the significance will be discussed in more detail in the subsequent analysis.

3.2. Mixed-ANOVA results

The Mixed-ANOVA results (see Table 2) show the effects of GenAI feedback type (non-visualised Group A vs. visualised Group B) and time (Pre vs. Post) on students' writing experiences. Across positive emotions, the main effect of time was highly significant ($p < .001$), indicating that students' CON ($\eta^2_p = 0.35$), ENY ($\eta^2_p = 0.58$), and WTW ($\eta^2_p = 0.34$) significantly improved over time, suggesting that GenAI feedback played a key role in fostering greater engagement in the revision process. However, the main effect of Group was not significant for these factors, implying that both chatbots supported improvements in writing experience, regardless of visualisation. While positive writing emotions improved, the analysis also revealed significant reductions in negative emotions. ANX significantly decreased over time ($p < .001$, $\eta^2_p = 0.28$), and group ($p = 0.034 < 0.05$, $\eta^2_p = 0.08$). According to Post hoc comparisons (see Table 3), after revising the essay with the help of visulised feedback, Group B had significantly lower anxiety than Group A ($t = 2.203$, $p = .032$*). Meanwhile, boredom significantly decreased over time ($p = .003$*, $\eta^2_p = 0.140$), and there was also a significant main effect of Group ($p = 0.024$*, $\eta^2_p = 0.09$). Following the essay revision process with visualised feedback, Group B reported significantly lower boredom levels compared to Group A ($t = 2.542$, $p = .014$*), as presented in Table 3. This finding indicates that the use of red highlights for errors, green highlights for correct revisions, and visually structured tables made the feedback process more engaging and enjoyable while also alleviating the cognitive demands of self-correction.

The analysis on revision quality revealed significant main effects for Time across all variables, indicating these observed improvements over time were more likely driven by common factors across both groups, such as the practice effect brought by GenAI or the general benefits of receiving prompt feedback. Additionally, significant differences between the groups were observed for all variables except GA, suggesting that both groups improved similarly in grammatical accuracy, regardless of the feedback type. According to Table 3, post-hoc comparisons for AVE showed a significant difference ($t = -2.120$, $p = 0.038$*). This suggests that, despite similar scores between the two groups at the pre-intervention time point, visualised feedback led to a more significant improvement in overall revision quality in the final results rather than in the rate of improvement.

Moreover, the significant interaction effect of CC ($F = 5.33$, $p = 0.025$*, $\eta^2_p = 0.080$) in Table 2 suggests that the improvement in CC differed significantly between the two groups over time. The post-hoc comparisons presented in Table 4 further revealed that Group B outperformed Group A in the post-test (*Estimate* $= -0.600$, $SE = 0.180$, $t = -3.325$, $p = 0.009^{**}$), suggesting that visualised feedback led to better outcomes in CC. Additionally, Group B significantly outperformed Group A from pre-test to post-test (*Estimate* $= 1.433$, $SE = 0.169$, $t = 8.470$, $p < 0.001$), with no significant difference between the baselines of Group A and Group B ($p = 1.000$). This means that, after excluding the interference of initial ability differences, the intervention of visualised feedback directly led to a greater improvement in CC. Its dynamic, real-time representation of textual structure likely provided learners with clearer targets for revision.

Table 1
Descriptive data.

Group	Non-Visualised (Group A)				Visualised (Group B)			
Time	Pre		Post		Pre		Post	
Dependent Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
WTW	3.613	0.825	4.027	0.755	3.707	0.909	4.340	0.554
ANX	3.071	0.689	2.671	1.012	2.820	0.710	2.142	0.841
ANY	3.453	0.714	4.227	0.668	3.500	0.735	4.373	0.500
CON	3.500	0.692	4.025	0.826	3.608	0.672	4.175	0.562
Boredom	2.600	0.767	2.442	1.035	2.300	0.794	1.817	0.861
CL	N/A	N/A	3.173	0.823	N/A	N/A	2.527	0.933
TL	N/A	N/A	3.980	0.880	N/A	N/A	3.973	0.822
SA	N/A	N/A	4.347	0.575	N/A	N/A	4.233	0.561
TR	5.833	0.648	6.700	0.651	6.067	0.640	6.967	0.669
CC	5.800	0.610	6.633	0.809	5.900	0.607	7.233	0.568
LR	5.800	0.484	6.800	0.610	6.067	0.640	7.033	0.718
GRA	5.967	0.669	6.933	0.828	6.033	0.615	7.100	0.712
AVE	5.850	0.429	6.767	0.636	6.017	0.495	7.083	0.514

M: Mean, SD: Standard Deviation, WTW: Willingness to Write, ANX: Foreign Language Writing Anxiety, ENY: Foreign Language Writing Enjoyment, CON: Foreign Language Writing Confidence, Boredom: Foreign Language Writing Boredom.

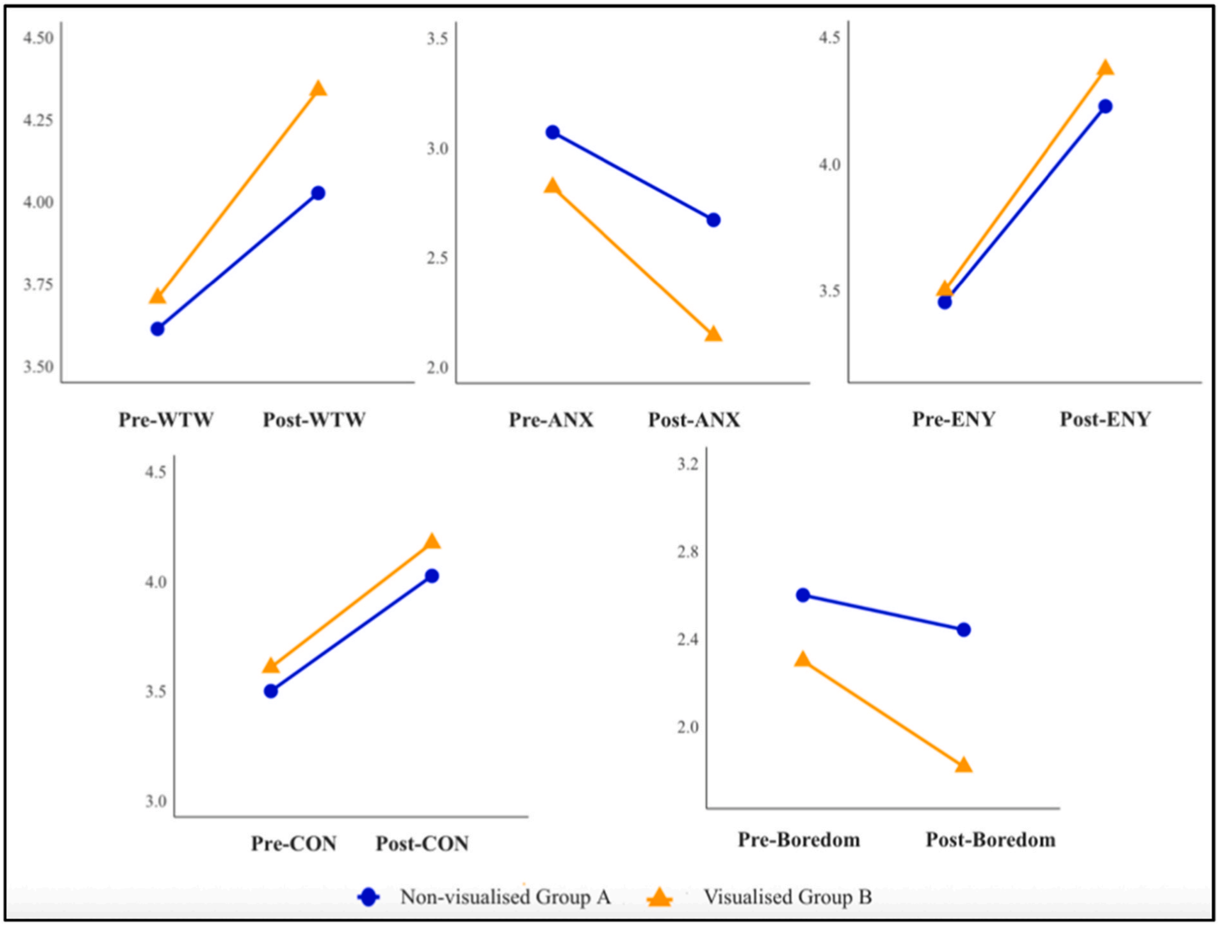

Fig. 7. The changes in WTW, ANX, ENY, CON, and Boredom over time are in two groups.

Such visual scaffolding actively facilitates writers' ability to perceive and address discourse-level issues.

3.3. Independent sample t-test result

As shown in Table 5, Group A exhibited a significantly higher CL than Group B ($t(58) = 2.847$, $p = 0.006 < 0.01$), indicating that visualised feedback effectively reduced cognitive demands during the revision process. The effect size ($d = 0.880$) suggests a medium to large effect, underscoring the practical significance of this difference. There was no significant difference in TL or SA between the two groups ($t(58) = 0.032$, $p = 0.974$; $t(58) = 0.773$, $p = 0.443$), suggesting that the feedback format had no measurable impact on students' perceived similarity to a human teacher or their overall satisfaction of the chatbots.

4. Discussion

This study extends the CATLM by demonstrating how positively designed visual elements in AI-mediated feedback can enhance coherence and cohesion in EFL writing, alleviate negative emotions, and shape learners' cognitive load. It offers new evidence for the dynamic interplay between cognitive and affective processes in GenAI-assisted language learning environments. The following sections will discuss these three aspects through the lens of CATLM.

4.1. Better revision quality with the help of visualised GenAI-feedback

The results of both descriptive analysis and mixed-ANOVA indicate that visualised GenAI feedback significantly enhanced learners' revision quality. Notably, the improvement in coherence and cohesion was statistically significant in terms of both the main effect of the group and the interaction effect with time. This suggests that the visual elements embedded in the feedback, such as colour-coded highlights, bold formatting, and structured tables, meaningfully supported learners' ability to revise at the discourse level. As highlighted in previous research, visual design can aid in information organisation, reduce cognitive effort, and enhance learning outcomes

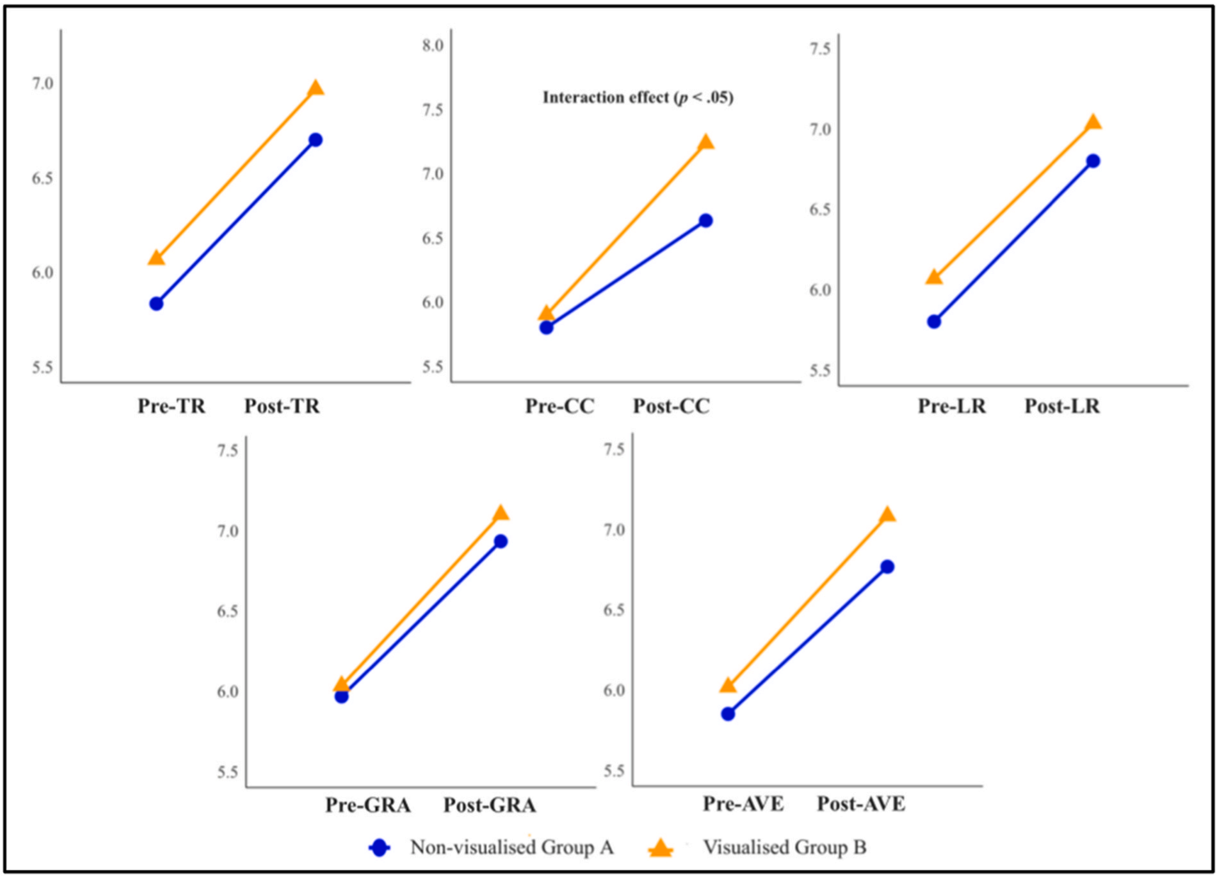

Fig. 8. The changes of TR, CC, LR, GRA, and AVE over time are in two groups.

(Mayer, 2005; Mayer & Estrella, 2014). This improvement can be attributed to the enhanced salience and accessibility of textual structural cues provided through visualised feedback. By clearly differentiating accurate from problematic areas (e.g., green highlights for correct revisions and red for errors), the system enabled learners to more efficiently and accurately identify revision targets. Moreover, the implementation of structured tables and segmented layouts likely contributed to a reduction in extraneous cognitive load (Liu et al., 2021), thereby enabling learners to allocate greater attentional resources to higher-level writing processes such as logical organisation, coherence, and argument development (Knight et al., 2020). These visual scaffolds appear to support learners in managing cognitive demands during revision, facilitating deeper engagement with discourse-level improvements (Barrot, 2023; Zhao, 2023).

Given that both groups started from comparable baseline levels, the statistically significant post-test differences in CC further underscore the impact of visualisation on improving global writing organisation, echoing previous research (Guo & Wang, 2023; Lu et al., 2024; Wei & Li, 2023). These findings are also in line with the CTML (Mayer, 2005) and CATLM (Moreno, 2006), which posits that multimodal representations facilitate deeper comprehension and improved retention. Such visual scaffolding functioned not only as an organisational aid but also as an intrinsic motivational cue, guiding learners toward more effective and engaging practices, ultimately contributing to learning outcomes (Hegarty, 2004; Li et al., 2025). The results reinforce the pedagogical value of incorporating visual elements into GenAI feedback systems to promote the revision process, particularly in areas requiring macro-level restructuring. Moreover, the findings suggest that chatbots designed in alignment with the CATLM and informed by the IELTS writing feedback process possess substantial practical value. By simulating teacher-like feedback, GenAI facilitates personalised and self-directed writing revision, while also helping to reduce teachers' workload and offering promising implications for enhancing the well-being of teachers.

4.2. Less negative EFL writing emotions after receiving visualised GenAI-feedback

Beyond enhancing learners' revision quality, the findings reveal that visualised GenAI feedback had a significant impact on reducing negative emotional responses commonly experienced during EFL writing tasks, particularly anxiety and boredom. The mixed-ANOVA results showed not only significant main effects of time for both variables ($p < 0.01$) but also a significant group effect for anxiety ($p = 0.034$, $\eta^2_p = 0.08$) and boredom ($p = 0.024$, $\eta^2_p = 0.09$). These results underscore the greater emotional benefit of

Table 2
Mixed-ANOVA results.

Dependent Variable	Effect	*F*-value	*p*-value	η^2 partial
WTW	Group	1.360	0.249	0.020
	Time	29.490	***	0.340
	Group × Time	1.300	0.258	0.020
ANX	Group	4.710	0.034*	0.080
	Time	22.630	***	0.280
	Group × Time	1.500	0.225	0.030
ENY	Group	0.460	0.503	0.010
	Time	79.010	***	0.580
	Group × Time	0.290	0.591	0.010
CON	Group	0.740	0.393	0.010
	Time	30.920	***	0.350
	Group × Time	0.050	0.833	0.001
Boring	Group	5.400	0.024*	0.090
	Time	9.380	0.003**	0.140
	Group × Time	2.410	0.126	0.040
TR	Group	4.680	0.035*	0.070
	Time	52.050	***	0.470
	Group × Time	0.020	0.892	< .001
CC	Group	7.250	0.009**	0.110
	Time	100.020	***	0.630
	Group × Time	5.330	0.025*	0.080
LR	Group	4.360	0.041*	0.070
	Time	86.360	***	0.600
	Group × Time	0.020	0.875	< .001
GA	Group	0.760	0.386	0.010
	Time	65.600	***	0.530
	Group × Time	0.160	0.692	0.003
AVE	Group	5.680	0.020*	0.090
	Time	122.620	***	0.680
	Group × Time	0.700	0.406	0.010

WTW: Willingness to Write, ANX: Foreign Language Writing Anxiety, ENY: Foreign Language Writing Enjoyment, CON: Foreign Language Writing Confidence, Boredom: Foreign Language Writing Boredom, TR: Task Response, CC: Coherence and Cohesion, LR: Lexical Resource, GA: Grammatical Accuracy, AVE: Average Score

* $p < 0.05$
** $p < 0.01$,
*** $p < 0.001$,

Table 3
Post-hoc comparisons results (pairwise ~ Group | Time).

Dependent Variable	Time	Estimate	*SE*	*t*	*p*
WTW	Post	−0.313	0.171	−1.833	0.072
	Pre	−0.093	0.224	−0.416	0.679
ANX	Post	0.529	0.240	2.203	0.032*
	Pre	0.251	0.181	1.391	0.170
ENY	Post	−0.147	0.152	−0.963	0.340
	Pre	−0.047	0.187	−0.249	0.804
CON	Post	−0.150	0.182	−0.822	0.414
	Pre	−0.108	0.176	−0.615	0.541
Boredom	Post	0.625	0.246	2.542	0.014*
	Pre	0.300	0.202	1.488	0.142
TR	Post	−0.267	0.170	−1.565	0.123
	Pre	−0.233	0.166	−1.404	0.166
LR	Post	−0.233	0.172	−1.356	0.180
	Pre	−0.267	0.146	−1.821	0.074
GA	Post	−0.167	0.199	−0.836	0.406
	Pre	−0.067	0.166	−0.402	0.689
AVE	Post	−0.317	0.149	−2.120	0.038*
	Pre	−0.167	0.120	−1.394	0.169

Post = Group A post vs. Group B post, Pre = Group A pre vs. Group B pre;
* $p < 0.05$

visualised feedback over non-visualised alternatives, with post hoc comparisons confirming that learners in the visualised condition reported significantly lower post-test anxiety and boredom levels than those who received standard text-only feedback.

These findings are consistent with the CATLM (Moreno, 2006). As learners' affective states dynamically interact with cognitive

Table 4
Post-hoc comparisons results (pairwise ~ Group * Time).

Dependent Variable	Pair	Estimate	*SE*	*t*	*p*
CC	A post - B post	−0.600	0.180	−3.325	0.009**
	A post - A Pre	0.833	0.153	5.440	***
	A post - B Pre	0.733	0.169	4.333	***
	B post - A Pre	1.433	0.169	8.470	***
	B post - B Pre	1.333	0.153	8.704	***
	A Pre - B Pre	−0.100	0.157	−0.636	1.000

* $p < 0.05$
** $p < 0.01$,
*** $p < 0.001$,

Table 5
Independent sample *t*-test.

Variable	*t*	df	*p*	Cohen's *d*	*95 % CI*
					lower upper
CL	2.847	58.000	0.006**	0.880	[0.192, 1.101]
TL	0.032	58.000	0.974	0.800	[−0.407, 0.421]
SA	0.773	58.000	0.443	0.568	[−0.180, 0.407]

CL: Cognitive Load, TL: Teacher Likeness, SA: Satisfaction
** $p < 0.01$,

processing, the design of multimedia environments, through coloured visual features in this study, can positively shape both emotion and learning outcomes (Plass et al., 2014; Tan et al., 2021). As Li and Han (2022) point out, boredom is typically considered a relatively stable and resistant emotional state in the context of foreign language learning, often more difficult to alleviate than other transient negative emotions. Wang et al. (2025) further revealed that university-level EFL learners who reported high levels of boredom were also more likely to resist engaging with AI-mediated learning environments, perceiving them as impersonal or cognitively taxing. In response to Wang's (2023) call for enhancing HCI quality to alleviate boredom in online instruction, the current study demonstrates that integrating visual reinforcements into GenAI-generated feedback can serve as cognitive organisers and function as emotionally responsive design elements. The embedding of such visual cues into self-revision tasks enhanced learners' sense of clarity, direction, and control, thereby contributing to a reduction in perceived monotony and disengagement during the revision process.

More importantly, these findings offer practical implications for supporting learners who are prone to disengagement or scepticism toward GenAI tools. Graphical and coloured representations can make complex structures more intuitive, accessible, and easier to understand, promoting comprehension, enjoyment, and memory retention and offering a psychologically accessible interface for sustained engagement (Hegarty, 2004; Li et al., 2025; Wang & Zou, 2025). This suggests that GenAI tools, when combined with principles of HCI design and CATLM (Moreno, 2006), can be particularly effective in supporting emotionally vulnerable or less motivated EFL learners, turning a passive revision task into an interactive and emotionally affirming experience.

4.3. Lower cognitive load after engaging with visualised GenAI-feedback

While the visualised GenAI feedback demonstrated measurable cognitive and emotional benefits, learners' perceptions of teacher likeness and satisfaction remained comparably high across both groups. This indicates that the overall quality of HCI was positively evaluated, regardless of the feedback's visual format. In other words, learners did not equate visual elements with greater "human-likeness" or satisfaction, suggesting that these affective judgments are shaped by broader communicative and relational cues as well as the revision flow beyond surface-level interface design (see Figs. 3 & 4). Perceived teacher likeness in AI-mediated environments is often linked to how well the system replicates human qualities such as clarity, relevance and responsiveness (Tram et al., 2024). In this study, both chatbots were designed to simulate teacher-like revision processes and flow. The consistently high TL scores suggest that learners interpreted the feedback as sufficiently didactic and supportive, possibly due to the use of clear language, structured commentary, and goal-oriented suggestions.

Similarly, learner satisfaction appeared to be driven not solely by aesthetics or formatting but by task relevance, perceived utility, and learning gain (Ca, Lin & Yu, 2023; Huang & Zou, 2024). Prior studies in computer-assisted language learning have emphasised that learner satisfaction is often rooted in perceived effectiveness and usefulness (Lin & Yu, 2023). In this case, both chatbots provided timely, task-specific, and grammatically accurate feedback, which likely contributed to learners' overall positive evaluation.

In terms of CL, the significantly lower cognitive load reported by learners receiving visualised GenAI feedback ($p = 0.006^{**}$, $d = 0.880$) suggests that the use of visual stimuli helped reduce extraneous mental effort. This reduction likely enabled learners to allocate more cognitive resources to discourse-level improvements, which may partly explain the greater enhancement in coherence and cohesion observed in this group. The finding aligns with CLT (Sweller et al., 2011), which emphasises the importance of instructional design in optimising mental processing during complex tasks. Unlike the conventional plain-text display modes of most GenAI tools in the market, learners with lower cognitive load can focus more on higher-order writing components such as discourse organisation. This may enhance their sense of control and learning efficiency during the revision process (Plass et al., 2014; Plass & Kaplan, 2016). The reduced risk of information overload may also partially explain why Group B exhibited lower negative emotions following the visualised feedback, as CATLM (Moreno, 2006) posits that emotionally supportive design can ease cognitive processing and foster more positive affective experiences during learning. The visual elements also offer a potential solution to the high cognitive load experienced by EFL learners during prompt design, as reported by Woo et al. (2024).

The findings of this study carry several pedagogical and design-oriented implications. For language teachers and learners, the integration of GenAI tools that offer visualised feedback holds promise in alleviating writing-related negative emotions, particularly in self-directed revision contexts. EFL teachers can adopt similar visualised GenAI feedback to support students in independently revising their writing, thereby enhancing the learning experience and potentially reducing teachers' workload. Embedding such tools into process-oriented writing instruction is particularly valuable in resource-limited instructional contexts. From a development perspective, thoughtful visual formatting, including the use of colour, structural clarity, and textual emphasis, appears to play an important role in improving the clarity and usability of feedback. Future GenAI writing platforms should therefore align feedback presentation with CLT (Sweller et al., 2011) and CATLM (Moreno, 2006): calibrate visual salience to cognitive capacity, embed affective cues to encourage persistence, and offer adaptable formats to accommodate learner diversity. Such theoretically grounded design will maximise the long-term educational impact of GenAI-mediated feedback systems.

5. Conclusions and limitations

This study explored how two feedback formats, visualised and plain-text, delivered by GenAI-based chatbots with identical settings and interaction modes, influenced EFL learners' revision performance, emotional responses, and interaction experiences. Both feedback formats proved helpful in guiding learners through revision. More importantly, the visualised feedback was particularly effective in improving coherence and cohesion, which suggests its value in supporting global aspects of writing quality. The visualised feedback also contributed to reducing negative emotions such as anxiety and boredom. This finding highlights the importance of positive emotional design in technology-mediated learning environments and further validates the relevance of CATLM (Moreno, 2006) in the current era of widespread GenAI adoption. Both groups reported a strong sense of teacher likeness during their interaction with the chatbot, suggesting a generally smooth and acceptable GenAI-powered revision process. Learners who received visualised feedback also reported a lower cognitive load, reflecting a more supportive cognitive learning experience.

Despite the study's innovative contribution to the growing body of research on GenAI-mediated writing feedback, several limitations should be acknowledged. First, the relatively small sample size may limit the extent to which the findings can be generalised to broader EFL learner populations. Second, the study was constrained to a single round of revision due to time and logistical considerations. As writing development is often a cumulative process, future research is encouraged to adopt longitudinal or multi-stage designs to capture the sustained effects of visualised feedback. Third, the absence of qualitative data constrained a more nuanced understanding of learners' emotional experiences and engagement with the GenAI feedback. Future research employing mixed-methods approaches may yield richer insights into learners' perceptions and interactions with GenAI tools. Finally, the visual design elements (e.g., colour-coding) were not culturally adapted. Given potential cultural variability in interpreting visual cues, future studies should consider culturally responsive design to enhance the cross-contextual applicability of GenAI feedback.

Funding Information

This study is supported by the TDF 202324S2–73 and RDF-24–01–058 at Xi'an Jiaotong-Liverpool University, China. The study also received funding from the Reform of Foreign Language Teaching under the Background of High-Quality Development in Universities of Jiangsu Province (Key project): (2024WYJG018).

CRediT authorship contribution statement

Bin Zou: Writing – original draft & review, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Chenghao Wang:** Conceptualization, Methodology, Software, Resources, Investigation, Data curation, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Huimin He:** Writing – original draft, Project administration. **Congxin Li:** Visualisation, Data curation, Methodology. **Erick Purwanto:** Visualisation, Data curation, Funding acquisition. **Ping Wang:** Writing – Funding acquisition.

Acknowledgements

This research was supported by the Teaching Development Fund (TDF 202324S2–73) and the Research Development Fund (RDF-24–01–058) at Xi'an Jiaotong-Liverpool University, Suzhou, China. The authors also gratefully acknowledge the funding from the Jiangsu Provincial Key Project for the Reform of Foreign Language Teaching in the Context of High-Quality University Development (2024WYJG018).

References

Abdullayeva, M., & Musayeva, Z. M. (2023). The impact of Chat GPT on student's writing skills: an exploration of AI-assisted writing tools. In *International Conference of Education, Research and Innovation, 1* pp. 61–66).

Alsofyani, A. H., & Barzanji, A. M. (2025). The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: a mixed-methods study. *Journal of Educational Computing Research, 63*(2), 431–463. https://doi.org/10.1177/07356331241307297

Barrot, J. S. (2023). Using ChatGPT for second language writing: pitfalls and potentials. *Assessing Writing: An International Journal, 57*.

Botes, Elouise, Greiff, Elouise, & Dewaele, Jean-Marc (2020). The foreign language classroom anxiety scale and academic achievement: an overview of the prevailing literature and a meta-analysis. *The Journal for the Psychology of Language Learning, 2*, 26–56.

Boudouaia, A., Mouas, S., & Kouider, B. (2024). A study on ChatGPT-4 as an innovative approach to enhancing English as a foreign language writing learning. *Journal of Educational Computing Research, 62*(6), 1509–1537. https://doi.org/10.1177/07356331241247465

Brom, C., Starkova, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colours during multimedia learning. *Educational Research Review, 25*, 100–119. https://doi.org/10.1016/j.edurev.2018.09.004

Brown, H. D. (1994). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. Prentice Hall Regents.

British Council. (2023). IELTS writing band descriptors (public version) [PDF]. ⟨https://takeielts.britishcouncil.org/sites/default/files/ielts_writing_band_descriptors.pdf⟩.

Cai, Q., Lin, Y., & Yu, Z. (2023). Factors influencing learner attitudes towards ChatGPT-assisted language learning in higher education. *International Journal of Human-Computer Interaction*, 1–15. https://doi.org/10.1080/10447318.2023.2261725

Cong-Lem, N., Soyoof, A., & Tsering, D. (2025). A systematic review of the limitations and associated opportunities of ChatGPT. *International Journal of Human-Computer Interaction, 41*(7), 3851–3866.

Derakhshan, A., Doliński, D., Wang, Y., & Fan, Z. (2025). An introduction to the special issue on the applications of positive psychology in educational environments: Teachers' and students' psychological states in the spotlight. *Learning and Motivation.* , Article 102116. https://doi.org/10.1016/j.lmot.2025.102116

Dewaele, J. M., & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety: the right and left feet of FL learning? In P. D. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive Psychology in SLA. Multilingual Matters*.

Dong, L., Liu, M., & Yang, F. (2022). The relationship between foreign language classroom anxiety, enjoyment, and expectancy-value motivation and their predictive effects on Chinese high school students' self-rated foreign language proficiency. *Frontiers in Psychology, 13*.

Du, J., & Alm, A. (2024). The impact of ChatGPT on English for academic purposes (EAP) students' language learning experience: a self-determination theory perspective. *Education Sciences, 14*(7), 726. https://doi.org/10.3390/educsci14070726

Du, J., & Daniel, B. K. (2024). Transforming language education: a systematic review of AI-powered chatbots in english as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, Article 100230. https://doi.org/10.1016/j.caeai.2024.100230

Feng, H., Li, K., & Zhang, J. (2025). What does AI bring to second language writing? A systematic review (2014-2024). *Language Learning Technology, 29*(1), 1–14.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Ghafouri, M. (2024). ChatGPT: the catalyst for teacher-student rapport and grit development in L2 class. *System, 120*.

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82*(3), 525–538. https://doi.org/10.1037/0022-0663.82.3.525

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

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI Chatbot-based learning: alleviating students' anxiety in English Writing Classroom. *Bulletin of Social Informatics Theory and Application, 7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

Hegarty, M. (2004). Dynamic visualisations and learning: getting to the difficult questions. *Learning and Instruction, 14*(3), 343–351.

Horwitz, E. (2017). On the misreading of Horwitz, Horwitz and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New Insights Into Language Anxiety: Theory, Research and Educational Implications* (pp. 31–47). Bristol: Multilingual Matters.

Huang, J., & Mizumoto, A. (2025). The effects of generative AI usage in EFL classrooms on the L2 motivational self system. *Education and Information Technologies, 30*(5), 6435–6454. https://doi.org/10.1007/s10639-024-13071-6

Huang, J., & Tan, M. (2023). The role of ChatGPT in scientific communication: writing better scientific review articles. *American Journal of Cancer Research, 13*(4), 1148.

Huang, F., Wang, Y., & Zhang, H. (2024). Modelling generative AI acceptance, perceived teachers' enthusiasm and self-efficacy to English as a Foreign Language Learners' Well-Being in the Digital Era. *European Journal of Education, 59*(4), Article e12770. https://doi.org/10.1111/ejed.12770

Huang, F., & Zou, B. (2024). English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI and innovativeness. *Computers in Human Behavior, 159*, Article 108355. https://doi.org/10.1016/j.chb.2024.108355

Huang, Z., Hu, X., & Jin, H. (2024). The predictive effects of L2 writing anxiety on motivational regulation strategies: a person-centered approach. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12677

Hwang, W.-Y., Nurtantyana, R., Purba, S. W. D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). AI and recognition technologies to facilitate English as a foreign language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research, 61*(5), 1008–1035. https://doi.org/10.1177/07356331221137253

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Knight, S., Shibani, A., Abel, S., Gibson, A., & Ryan, P. (2020). AcaWriter: a learning analytics tool for formative feedback on academic writing. *Journal of Writing Research, 12*(2), 141–186. https://doi.org/10.17239/jowr-2020.12.01.06

Li, M., Wang, D., Purwanto, E., Selig, T., Zhang, Q., & Liang, H.-N. (2025). VisualCodeMOOC: a course platform for algorithms and data structures integrating a conversational agent for enhanced learning through dynamic visualisations. *SoftwareX, 30*, Article 102072. https://doi.org/10.1016/j.softx.2025.102072

Li, C., & Han, Y. (2022). The predictive effects of foreign language enjoyment, anxiety, and boredom on learning outcomes in online English classrooms. *Modern Foreign Languages, 45*(2), 207–219.

Li, J., & Wang, J. (2024). A measure of EFL argumentative writing cognitive load: Scale development and validation. *Journal of Second Language Writing, 63*, Article 101095. https://doi.org/10.1016/j.jslw.2024.101095

Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning: An International Journal, 35*(4), 605–634. https://doi.org/10.1080/09588221.2020.1743323

Liu, C., Hou, J., Tu, Y., Wang, Y., & Hwang, G. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments, 31*(9), 5614–5632. https://doi.org/10.1080/10494820.2021.2012812

Liu, M., Zhang, L. J., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers Education, 211*, Article 104977. https://doi.org/10.1016/j.compedu.2024.104977

Lu, Q., Yao, Y., Xiao, L., Yuan, M., Wang, J., & Zhu, X. (2024). Can ChatGPT effectively complement teacher assessment of undergraduate students' academic writing? *Assessment Evaluation in Higher Education, 49*(5), 616–633.

Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learners' argumentative writing. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13488-7

Lin, Y., & Yu, Z. (2023). A meta-analysis evaluating the effectiveness of instructional video technologies. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-023-09669-3

Macklem, G. L. (2015). *Boredom in the Classroom: Addressing Student Motivation, Self-Regulation, and Engagement in Learning*. Cham: Springer.

MacIntyre, P. (2017). 2. An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J. Dewaele (Eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 11–30). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781783097722-003.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualising willingness to communicate in an L2: a situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., & Gregersen, T. (2012a). In S. Mercer, S. Ryan, & M. Williams (Eds.), *Affect: The Role of Language Anxiety and Other Emotions in Language Learning. In Language Learning Psychology: Research, Theory and Pedagogy* (pp. 103–118). Basingstoke: Palgrave.

MacIntyre, P. D., & Gregersen, T. (2012b). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching, 2*(2), 193–213.

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. *Smart Learning Environments, 11*(1), 9. https://doi.org/10.1186/s40561-024-00295-9

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004.

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12–18. https://doi.org/10.1016/j.learninstruc.2014.02.004

Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia, 15*(1), 1–26.

Moreno, R. (2006). Does the modality principle apply to different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning, 22*(3), 149–158. https://doi.org/10.1111/j.1365-2729.2006.00170.x

Moreno, R., Mayer, R. E., Spires, H., & Lester, J. (2001). The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*(2), 177–213. https://doi.org/10.1207/S1532690XCI1902_02

Münchow, H., & Bannert, M. (2019). Feeling good, learning better? Effectivity of an emotional design procedure in multimedia learning. *Educational Psychology, 39*(4), 530–549. https://doi.org/10.1080/01443410.2018.1524852

Nie, J., Yuan, Y., Chao, X., Li, Y., & Lv, L. (2024). In smart classroom: investigating the relationship between human-computer interaction, cognitive load and academic emotion. *International Journal of Human-Computer Interaction, 40*(13), 3528–3538.

Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: an eye-tracking study on the use of anthropomorphisms. *Computers Education, 86*, 30–42. https://doi.org/10.1016/j.compedu.2015.02.016

Pekrun, R., & Perry, P. P. (2014). Control-value theory of achievement emotions. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 120–141). Routledge.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and colour on affect and learning. *Learning and Instruction, 29*, 128–140.

Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In M. Pekrun, L. Linnenbrink-Garcia, & H. M. Sinatra (Eds.), *Emotions, Technology, Design, and Learning* (pp. 131–161). Academic.

Prior, M. (2019). Elephants in the room: an "affective turn," or just feeling our way? *The Modern Language Journal, 103*(2), 516–527.

Rafiee, M., & Abbasian-Naghneh, S. (2020). Willingness to Write (WTW): development of a model in EFL writing classrooms. *Cogent Education, 7*(1), Article 1847710. https://doi.org/10.1080/2331186X.2020.1847710

Shao, K., Nicholson, L. J., Kutuk, G., & Lei, F. (2020). Emotions and instructed language learning: proposing a second language emotions and positive psychology model. *Frontiers in Psychology, 11*.

Sharrad, A., & Faravani, A. (2021). Delving into the relationships among Iraqi EFL learners' willingness to write, personality types and writing ability. *KURMANJ, 3*(1), 23–29.

Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning. *Journal of Computer Assisted Learning, 35*(4), 555–568. https://doi.org/10.1111/jcal.12359

Sweller, J. (2008). Human cognitive architecture. *Handbook of Research on Educational Communications and Technology* (pp. 369–381). Routledge.

Sweller, J., Ayres, P., & Kalyuga, S., 2011. The Guidance Fading Effect. Cognitive load theory. Explorations in the Learning Sciences, Instructional Systems and Performance Technologies (Vol. 1, pp. 171–182). Springer..

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

Tan, J., Mao, J., Jiang, Y., & Gao, M. (2021). The influence of Academic emotions on Learning effects: a systematic review. *International Journal of Environmental Research and Public Health, 18*(18). https://doi.org/10.3390/ijerph18189678

Tanielian, A. R. (2017). Foreign language anxiety among first-year Saudi University students. *International Education Journal: Comparative Perspectives, 16*(2), 116–130.

Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence, 7*.

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: a multi-methods study. *System, 127*, Article 103528. https://doi.org/10.1016/j.system.2024.103528

Tran, T. T. T. (2025). Enhancing EFL writing revision practices: the impact of AI- and teacher-generated feedback and their sequences. *Education Sciences, 15*(2), 232. https://doi.org/10.3390/educsci15020232

Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2011). Emotional design in multimedia learning. *Journal of Educational Psychology, 104*(2), 485–498. https://doi.org/10.1037/a0026609

Wang, C., Li, X., & Zou, B. (2025). Revisiting integrated model of technology acceptance among the generative AI-powered foreign language speaking practice: through the lens of positive psychology and intrinsic motivation. *European Journal of Education, 60*(1).

Wang, C., Wang, Y., & Zou, B. (2024). Revolutionising EFL pedagogy: Innovative strategies for integrating GAI (ChatGPT) into language teaching. *Journal of Language Teaching, 4*(1), 1–7. https://doi.org/10.54475/jlt.2024.004

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: an analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System, 127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Wang, C., & Zou, B. (2025). D-ID Studio: empowering language teaching with AI avatars. *TESOL Journal, 16*(2), Article e70034. https://doi.org/10.1002/tesj.70034

Wang, C., & Li, X. (2025). Software review: empowering language education with D-ID creative reality studio's multimodal capabilities. *International Journal of Computer-Assisted Language Learning and Teaching, 15*(1), 1–11. https://doi.org/10.4018/IJCALLT.368218

Wang, M., & Wang, Y. (2024). A structural equation modelling approach in examining EFL students' foreign language enjoyment, trait emotional intelligence, and classroom climate. *Learning and Motivation, 86*, Article 101981. https://doi.org/10.1016/j.lmot.2024.101981

Wang, Y. (2023). Probing the boredom of online instruction among Chinese English language teachers during the COVID-19 pandemic. *Current Psychology, 43*, 12144–12158. https://doi.org/10.1007/s12144-023-04828-7

Wang, Y., Wang, Y., Pan, Z., & Ortega-Martín, J. L. (2024). The predicting role of EFL students' achievement emotions and technological self-efficacy in their technology acceptance. *The Asia-Pacific Education Researcher, 33*(4), 771–782. https://doi.org/10.1007/s40299-023-00750-0

Wang, Y., & Xue, L. (2024). Using AI-driven chatbots to foster Chinese EFL students' academic engagement: an intervention study. *Computers in Human Behavior, 159*, Article 108353. https://doi.org/10.1016/j.chb.2024.108353

Wang, Y., Derakhshan, A., & Solhi, M. (2025). Dispositions towards loving pedagogy, emotion regulation, and self-efficacy as predictors of EFL teachers' work engagement: a multinational study. *Language Teaching Research*. https://doi.org/10.1177/13621688251353132

Wang, H., Peng, A., & Patterson, M. M. (2021). The roles of class social climate, language mindset, and emotions in predicting willingness to communicate in a foreign language. *System, 99*, Article 102529.

Wang, Z. (2023). Psychological factors and production behaviours of Chinese undergraduate EFL learners. *PLoS One, 18*(7), Article e0288014. https://doi.org/10.1371/journal.pone.0288014

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, Article 1261955. https://doi.org/10.3389/fpsyg.2023.1261955

Wei, S., & Li, L. (2023). Artificial intelligence-assisted second language writing feedback: a case study of ChatGPT. *Foreign Languages in China, 20*(03), 33–40.

Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of emotional designs in multimedia learning: a replication and extension study. *Educational Psychology Review, 33*(2), 357–385. https://doi.org/10.1007/s10648-02009545x

Woo, D. J., Wang, D., Guo, K., & Zhang, Y. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies, 29*, 24963–24990. https://doi.org/10.1007/s10639-024-12819-4

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: experience and perceptions. *Languages, 8*(3), 212–231. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on Learners in an L2 Writing Practicum: An Exploratory Investigation. *Education and Information Technologies, 28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zhao, X. (2023). Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers. *RELC Journal: A Journal of Language Teaching and Research, 54*(3), 890–894. https://doi.org/10.1177/00336882221094089

Zou, B., Wang, C., Yan, Y., Du, X., & Ji, Y. (2024). Exploring English as a Foreign Language Learners' Adoption and Utilisation of ChatGPT for Speaking Practice Through an Extended Technology Acceptance Model. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12658

Bin Zou received his PhD degree in TESOL and Computer Technology from the University of Bristol, UK. He is a Senior Associate Professor and PhD supervisor at the Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University. His research interests include Computer-Assisted Language Learning (CALL), AI, EAP, and ELT. He is the Founding Editor and Co-Editor-in-Chief of two international journals: the International Journal of Computer-Assisted Language Learning and Teaching and the International Journal of EAP: Research and Practice.

Chenghao Wang is a second-year PhD student in the Department of Applied Linguistics at the University of Liverpool and Xi'an Jiaotong-Liverpool University. He also works as a researcher at ILABI Lab, Zhejiang University. His research interests include GenAI-assisted English speaking practice, EAP, ELT, and human–computer interaction in language education. His work has been published in journals such as System, International Journal of Human–Computer Interaction, and European Journal of Education, etc.

Huimin He is currently an English Language Lecturer at the English Language Centre, School of Languages at Xi'an Jiaotong-Liverpool University. She is also a fellow of the Higher Education Academy (HEA, the UK). Her research interests include second language writing, EAP pedagogy, discourse analysis and computer-assisted language learning.

Congxin Li is the Assistant Professor in the Department of Applied Linguistics, Xi'an-Jiaotong Liverpool University. She holds a PhD degree in Second Language Education at University of Cambridge. Her research interests include mobile-assisted language learning (MALL), computer-assisted language learning (CALL) and AI-assisted language learning with a focus on second language acquisition.

Dr. Erick Purwanto is a Teaching Fellow at Xi'an Jiaotong-Liverpool University (XJTLU) at the Computing Department of School of Advanced Technology. He received both Ph.D and M.Comp degree from National University of Singapore (NUS). His research interests are Machine Learning, Medical Imaging, and Technology Enhanced Learning.

Ping Wang joined Xi'an Jiaotong Liverpool University (XJTLU) in June 2023 as the Head of the Higher Education Division. He holds a BA, an MA in applied linguistics and an Ed.D in language education. Before he joined XJTLU, Professor Wang was a professor (level three) at Jiangsu University of Science and Technology. Dr. Ping Wang is Professor of language education. His main areas of interest are Translanguaging and Content and Language Integrated Learning.