Online EFL business writing with GenAI-generated templates: Students' performance and perceptions

Shu-Chiao Tsai
Center of Liberal Arts, National Kaohsiung University of Science and Technology, Taiwan, Republic of China

Generative artificial intelligence (GenAI) has been increasingly used to facilitate and evaluate EFL academic writing, but few studies have been conducted on business writing. Furthermore, little empirical research has evaluated the quality of the GenAI-generated templates. An approach using English GenAI-generated templates as feedback (GenAI-GETF) was implemented in two business contexts for 117 non-English major English as a Foreign Language (EFL) students in two groups, respectively. A pre-test and post-test design was conducted to assess students' writing performance, and a questionnaire survey was administered to investigate their perceptions. Results indicate that the GenAI-generated templates provided more enriched and professional content than samples excerpted from textbooks. Students showed significantly better performance in the post-writing than in the pre-writing, such as writing more enriched content, using more high-order words, having a lower probability of writing errors and performing higher writing scores. They also showed high satisfaction with their learning through the GenAI-GETF approach. The use of AI-based tools combined with student-centred activities assigned in this study provided a dynamic, interactive and self-paced learning environment in which students were able to immediately receive multimodal input, scaffolding and personalised feedback to improve their content-based and linguistic knowledge of their post-writing. These findings provide practical solutions and implications for the concerns often encountered in using technology.

Implications for practice or policy:

- The GenAI-GETF approach with a five-step learning model should be implemented in EFL business writing.
- EFL students should make good use of the features of GenAI and free AI-based resources to promote comprehensive input and output and obtain instant and personalised feedback for achieving improved writing performance.
- EFL teachers should adopt technological resources to create a flexible and interactive learning environment where students' interaction and outcomes can be effectively enhanced.

Keywords: generative artificial intelligence (GenAI), template, EFL, writing, business, performance, perception

Introduction

Business English writing is an important skill for students of English as a foreign language (EFL) to effectively handle and communicate in various tasks for their future career, such as letters, emails, memoranda, reports and minutes. Teaching business English writing skills presents EFL teachers with several challenges in limited classroom time with a high number of students in a class, including preparing instructional materials, developing students' language proficiency, teaching writing strategies, assigning learning activities, providing timely and personalised feedback and enhancing students' learning motivation and interaction (Arifin, 2017; Baskara, 2023; Isma et al., 2024, Tahriri et al., 2016; Tsai, 2017). In addition, the potential influence of EFL students' first language (L1) is another concern because L1 and second language (L2) are naturally linked to each other in their mind (Druce, 2012; Weijen et al., 2009). Process-based writing emphasises teaching writing as a process that generally involves the steps of planning, drafting, revising and editing. It has been widely adopted in different writing genres, including business writing, to help EFL students develop their ideas, receive feedback and refine their writing to achieve a final product of improved quality (Dokchandra, 2018; Guo et al., 2021; Mai, 2025).

With the development of innovative technologies, online translanguaging tools and automated writing evaluation systems based on artificial intelligence (AI) have been successfully implemented in EFL writing for providing students with instant and personalised feedback, enhancing their comprehensive language acquisition and improving their writing performance (Dizon & Gayed, 2024; Liang, 2025; Loncar et al. 2021; Teng, 2024; Tsai, 2022). For example, the use of generative AI (GenAI) tools provides several advantages in EFL writing, such as facilitating writing drafting, editing and proofreading (Allen & Mizumoto, 2024; Harunasari, 2023; Li & Kim, 2024), improving syntactic and semantic quality (Song & Song, 2023; Teng, 2024), suggesting language usage and sentence structures for supporting students' arguments (Baskara, 2023; Xiao & Zhi, 2023), promoting learners' motivation and engagement (Nguyen, 2023; Pham et al., 2023; Song & Song, 2023) and holding great potential for self-regulated learning (Li & Kim, 2024). Lu and Zeng (2025) found that using GenAI-generated model texts could yield comparable positive effects to using model texts composed by human teachers in an argumentative writing task for high school students. Regardless of the many advantages provided by using technology, there is still a concern about its disadvantages and potential threats to academic integrity, such as the frequency of errors, the risk of providing misleading information, inaccurate feedback, unethical use of the technology, plagiarism, overreliance on the tool and superficial interactions and distraction in the learning process (Darkoah et al., 2024; Kayalı et al., 2023; Shadiev & Yang, 2020; Teng, 2024; Wilkinson et al., 2024).

El Harrath (2024) emphasised that the application and development of computer-assisted language learning (CALL) should be driven by integrating second language acquisition (SLA) methods into curricula, assessment and tasks. In addition, according to the research-backed principles in the SLA field, Nielson (2022) indicated four requirements for a successful online learning experience: input, output, interaction and feedback, echoing Chapelle's (1998) suggested principles for developing multimedia CALL by emphasising input saliency and feedback, providing interaction for comprehensive output and allowing learners to focus on communication and engagement in the target language. Isbell et al. (2022) reported that it is necessary to take a strictly exploratory technology-enhanced SLA research method and adopt appropriate linguistic tests, survey tools and various types of measurements to collect diverse and detailed participants' behavioural parameters and investigate and verify their engagement, outcome and attitudes of technology. Furthermore, most studies have been conducted on academic EFL writing and few on business writing, one of the most valuable skills in the workplace for students' future career, and little empirical research has evaluated the quality of the GenAI-generated templates nor investigated its effects as feedback (Lu & Zeng, 2025). As students are major GenAI end users, it is essential to investigate their experiences and perceptions with its implementation to understand and support their specific needs, concerns, expectations and challenges (Li & Kim, 2024; Wilkinson et al., 2024).

To address the gap in our understanding of their needs, I proposed a technology-enhanced SLA approach with process-based writing to investigate the practicality and effectiveness of using GenAI-generated English templates as feedback (GenAI-GETF) in two business writing tasks (inquiry and collection) respectively assigned to two groups of non-English major EFL students. Their writing performance was qualitatively and quantitatively measured by using online computational and human assessments. The research questions (RQs) of the study were:

- RQ1: How is the text comparison between the two GenAI-generated English templates of inquiry and collection and the corresponding samples excerpted from commercial textbooks?
- RQ2: Can the writing performance of non-English EFL students be improved through learning with the GenAI-GETF approach?
- RQ3: What are the students' perceptions towards the GenAI-GETF approach?

Method

This study was implemented in Modern Technical Issues, an elective course in the liberal arts education track at a technical university, and approved by the Human Subjects Ethics Committee at National Cheng Kung University in Taiwan. In the online curriculum presented before course selection and in the first three weeks of formal classes, students were informed that this study would be implemented after the midterm exam. In the beginning of the study, I introduced the procedure, and then students had to download a worksheet explaining the steps for completing the assigned pre- and post-writing tasks from my server. Each student had a computer on which they wrote their scripts and interacted with the GenAI-generated templates. Two real-life writing tasks in the business field were assigned to two groups of the students, respectively:

- Task 1: Students in the INQ (inquiry) group had to play the role of a buyer to write an English letter of inquiry after having met a supplier of digital equipment at a trade show.
- Task 2: Students in the COL (collection) group had to play the role of a seller to write the second letter of collection in English for an overdue payment after having not received a reply from the buyer to the first letter of collection.

The procedure of the study is shown in Figure 1. I played the role of supervising and observing students' learning and controlling the instructional schedule.

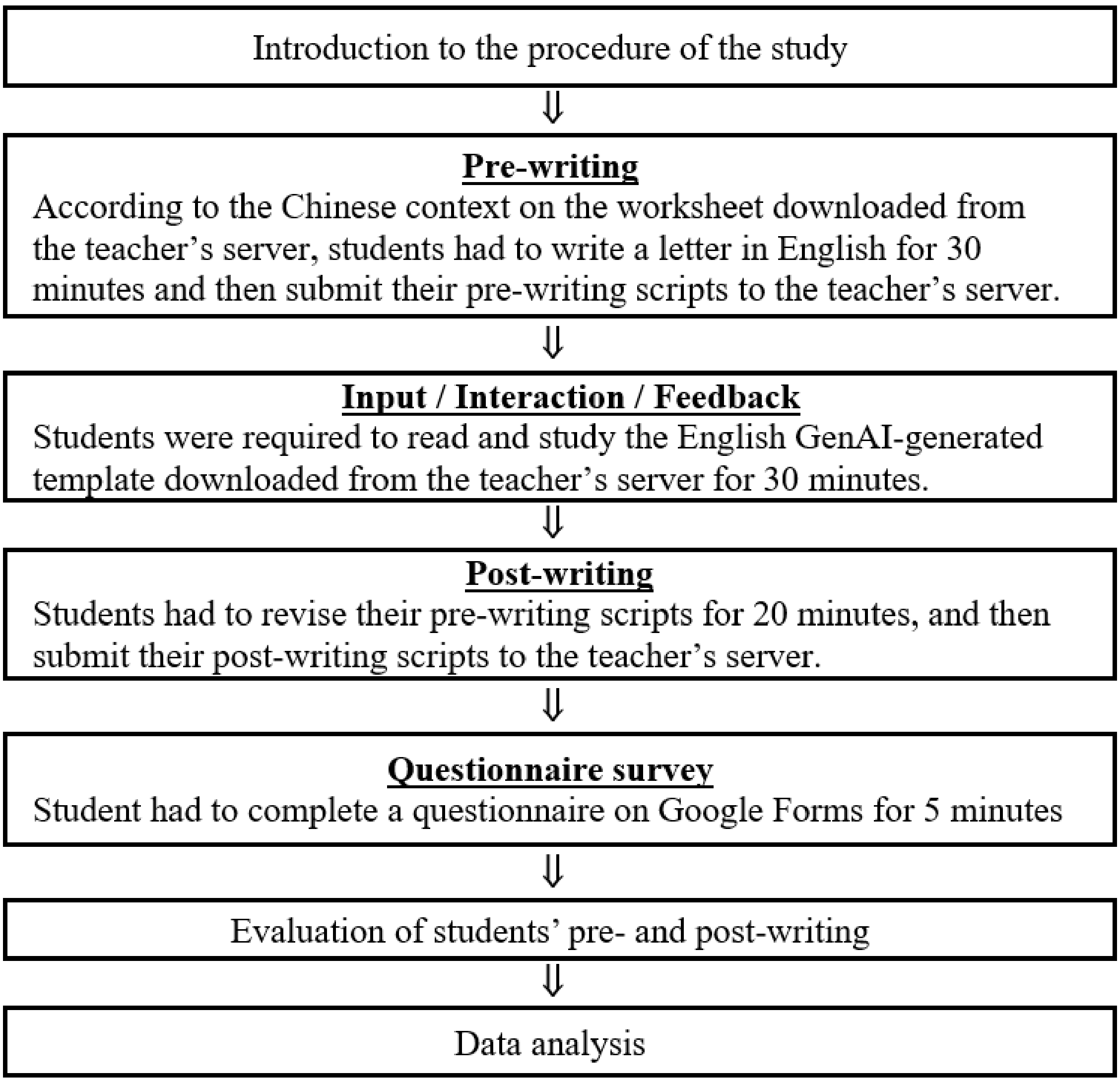

Figure 1. The procedure of the study

A five-step learning model of using the GenAI-GETF approach in business writing

Based on Nielson's (2022) requirements on the research-backed principles in the SLA field for a successful online learning experience, I proposed a five-step learning model of using the GenAI-GETF approach with process-based writing, as shown in Figure 2. ChatGPT was the main GenAI-based tool; the model is explained below.

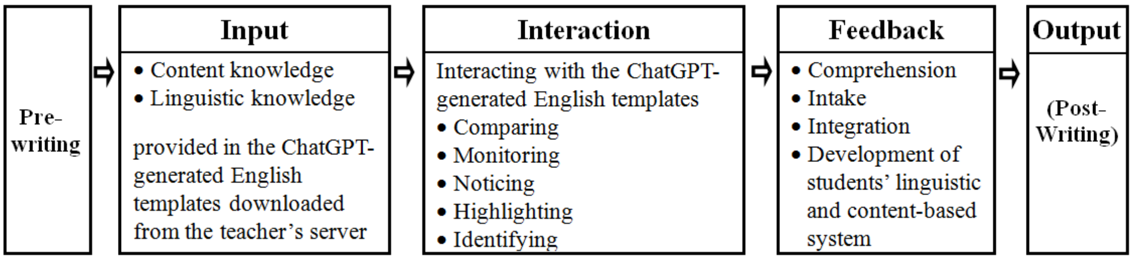

Figure 2. A five-step learning model of using the GenAI-GETF approach in business writing

1. Step 1 (Pre-writing): According to the Chinese context on the worksheet downloaded from the teacher's server, students had to write a letter in English for 30 minutes and then submit their pre-writing scripts to the server. They were able to use available online tools, but not ChatGPT, for checking vocabulary or grammar, improving semantic and syntactic errors or deficiencie, or translating complicated L1 sentences into L2. While conducting pre-writing, students were able not only to understand what would be expected of them to complete the writing tasks but also to think ahead about how to prepare the content and linguistic knowledge they would need in the subsequent steps.
2. Step 2 (Input): Students were required to read and study the English GenAI-generated template for 30 minutes. As comprehensive input in the target language is the potential starting point for acquiring aspects of the L2, students had to read the GenAI-generated English templates to identify patterns, concepts or principles regarding linguistic and content-based knowledge.
3. Step 3 (Interaction): By comparing their pre-writing scripts with the GenAI-generated templates, students were able to monitor, notice and identify their semantic and syntactic errors or deficiencies and focus on highlighting the content and language knowledge that they were unclear about, unfamiliar with or needed to know.
4. Step 4 (Feedback): As students received instant and personalised feedback through their interaction with important or unfamiliar aspects of the content and language knowledge by using AI-based tools, they were able to improve semantic and syntactic comprehension and had the potential for developing their content-based and linguistic system to become more prepared for the post-writing.
5. Step 5 (Output or post-writing): Students had to revise their pre-writing scripts for 20 minutes. Through a two-way interaction between students and the target content in the previous steps, students' writing could be improved to be syntactically well-formed and pragmatically appropriate toward the goal of the assigned tasks.

Participants

For convenience in sampling, non-English major students from four classes who took the optional course on Modern Technological Issues participated in this study. A total of 117 students (male: 106; female: 11), classified as Group ALL. The group included a subset of 59 students who had to complete a letter of inquiry, classified as INQ group and a subset of 58 students who had to write a letter of collection, classified as COL group. Their mean age was 21.0 years old, ranging from 19 to 22 years, and their English proficiency was determined by online simulated Test of English for International Communication (TOEIC) provided by the foreign language education centre of the university. Their TOEIC mean score was 515.4, equivalent to the B1 level (TOEIC score between 550 and 784). There was no significant difference in TOEIC scores between students in the two groups. The detailed background of the students is shown in Table 1.

Table 1
Background of the students

	Group		
	ALL	INQ	COL
No. of students	117	59	58
Gender			
Male	106	54	52
Female	11	5	6
Age: Mean	21.0	21.0	20.9
Simulated TOEIC score: Mean (SD)	515.4 (150.2)	520.9 (154.2)	509.8 (146.0)
No. of students who have used ChatGPT	115	58	57
No. of students who have taken the course in English business writing	8	5	3

Computational assessment

Four types of free online computational assessments were used to evaluate students' pre- and post-writing performance, as explained below (Tsai, 2019, 2022):

1. *Grammarly*: Grammarly is an AI-based automated writing evaluation tool; it can immediately provide corrective feedback and suggestions on grammar, vocabulary and mechanics and offer a score, ranging from 1 to 100, based on the highlighted corrections and offered suggestions that appear in the analysed text to represent writing quality (Barrot, 2023; Ding & Zou, 2024). In addition, the probability of writing errors is determined through the count of the errors identified by Grammarly in each script divided by the total words of the script.
2. *Vocabulary Profiler*: Vocabulary Profiler can classify the words of the analysed text in terms of the frequency with which such items appear in very large text corpora. The text words are divided into four categories (https://www.lextutor.ca/vp/eng/): the most frequent 1,000 words of English (K1), the second most frequent 1,000 words of English (K2), academic words list (AWL) and the remainder called "off-list words". For this study, the AWL and off-list words were classified as one group because they generally consisted of more advanced, difficult, and professional vocabulary related to lexical sophistication. Vocabulary Profiler also provides lexical density, by which the count of lexical words in the analysed text can be calculated. Lexical words, related to the use of different types of words, include nouns, adjectives, verbs and adverbs that are content words for giving meaning or information regarding what the text is about. Thus, the count of lexical words (LWs) can serve as a measure of the amount of information being conveyed in written texts. In addition, the count of different words (DWs) is provided.
3. *Document Readability*: Readability scores can give valuable insights into how easily a written text can be understood by readers (https://charactercalculator.com/flesch-reading-ease/). Flesch Readability Ease (FRE) scores indicate the understandability of a piece of content or a passage with a number ranging from 0 to 100. A text with a higher FRE score means its content is easier to read and understand.
4. *Self-developed programme*: I used a self-developed software programmed with JavaScript to investigate how the GenAI-generated English templates quantitatively influence students' post-writing. By comparison with the pre-writing scripts, the count of the increased words in the post-writing scripts can be determined; furthermore, the count of these increased words that are borrowed from the GenAI-generated English templates can also be measured. Accordingly, the percentage of the borrowed words can be calculated to survey the extent of students' reliance on the GenAI-generated templates.

Questionnaire

Students' perceptions of the GenAI-GETF approach in business writing were elicited through an online questionnaire on Google Forms consisting of two parts:

- Part 1 consisted of 16 items. The first 11 items (SQ1 to SQ11) focused on investigating whether students perceived GenAI as a potential tool for improving their business writing in content, vocabulary, grammar, sentence patterns, speed, ideas, material preparation and their willingness and recommendation for using GenAI. The other 5 items (SQ12 to SQ16) involved some issues regarding its impact on academic integrity, plagiarism, critical thinking and creativity.
- Part 2 involved students' opinions on the appropriateness and quality of the content in the GenAI-generated English templates. There were 11 items (TQ1 to SQ11) focusing on its grammar, expression, clarity, richness, coherence, length and structure.

The questionnaire, adapted from a previous study (Tsai, 2019, 2022), was modified and then reviewed by two experienced EFL teachers to ensure the content validity of the survey. In addition, the Kaiser–Meyer–Olkin (KMO) test, a statistical measure of factor analysis, was also used to evaluate the structural validity of the questionnaire. The questionnaire also included a freely given informed consent to participate. The questionnaire used a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). All returned questionnaires were analysed using IBM SPSS version 20.

Data analysis

Writing parameters in the GenAI-generated templates, students' pre- and post-scripts, and the samples excerpted from textbooks were determined by four types of online computational assessments. The professional points of the above scripts were examined and checked by two experienced EFL teachers. Students' responses to the questionnaires were collected and analysed by using IBM SPSS version 20 to calculate their means and standard deviations. In addition, the relationship among students' writing performance in the pre- and post- writing, their English proficiency, and their perceptions was analysed by using independent and paired sample *t* tests. The effect size (Cohen's *d*) was computed from the means and standard deviations of the items with statistically significant difference to measure the magnitude of the experimental effect. The Cohen's *d* value around 0.2 represents a small effect size, 0.5 a medium effect size and 0.8 a large effect size (Cohen, 1988; McLeod, 2023).

Results

RQ1: Text comparison

The GenAI-generated English template of each task incorporated with three corresponding samples of letters excerpted from textbooks were analysed (Holt & Sampson, 2004; Macintosh, 1994; Taylor, 2012), as explained below.

Task 1: Letter of inquiry

Based on the analysis by Vocabulary Profiler, the GenAI-generated template I provided presented 128 words, including 95 K1 words, 8 K2 words, 25 AWL and off-list words, with 69 lexical words and 94 different words. In addition, it had a writing score of 90, no writing errors and a FRE score of 36.8, corresponding to the level of college (FRE score between 30 and 50). According to the results evaluated by two experienced EFL teachers, 11 professional points mentioned in the task (IPPs) were checked and identified: a brief opening (IPP 1), the product to be inquired (IPP 2), asking for product catalogues (IPP 3), product specifications (IPP 4), product price (IPP 5), minimum order quantity (IPP 6), shipment (IPP 7), order procedure (IPP 8), asking for a reply (IPP 9), expressing gratitude (IPP 10) and looking for cooperation opportunities (IPP 11).

As for the samples of letters excerpted from textbooks, on average, each sample included about 105 words, including 85 K1 words, 5 K2 words, 15 AWL and off-list words, with 52 lexical words and 71 different words. All the three samples had no spelling or grammatical errors, and a writing mean score of 89, ranging from 88 to 90. In addition, the mean FRE score was 55.7, corresponding to the level of 10th to 12th grade (FRE score between 50 and 60). There were 5, 6, and 10 professional points in the three samples, respectively, which were almost mentioned in the GenAI-generated template. Two additional professional points, asking for after-sales service (IPP 12) and asking for a discount (IPP 13), were mentioned in one of the three samples but not in the GenAI-generated template.

Task 2: Letter of collection for the second round

The GenAI-generated template presented 158 words, including 127 K1 words, 11 K2 words, 20 AWL and off-list words, with 76 lexical words and 102 different words. It had no errors, a writing score of 85, and a FRE score of 60.5, corresponding to the level of 10th to 12th grade. Eight professional points mentioned in the task (CPPs) were provided: a brief opening (CPP 1), a reminder for not receiving the payment (CPP 2), no reply to the first letter of collection (CPP 3), overdue payment (CPP 4), asking for payment (CPP 5), an apology if the payment has been made before receiving this letter (CPP 6), asking for a quick reply (CPP 7) and salutation (CPP 8).

As for the three samples excerpted from textbooks, on average, each sample had 96 words, including 85 K1 words, 6 K2 words, 5 AWL and off-list words, with 45 lexical words and 63 different words. Based on the evaluation by Grammarly, all the three samples had no spelling or grammatical errors, and had a writing mean score of 89, ranging from 81 to 96. In addition, the mean FRE score was 71.2, corresponding to the level 7th grade (FRE score between 70 and 80). There were 5, 6, and 6 professional points mentioned in the three samples, respectively. Two of these professional points, identifying possible problems (CPP 9) and taking legal action (CPP 10), were respectively mentioned in two samples but not in the GenAI-generated template.

RQ2: Students' writing performance

The results from the Pearsons' analysis showed that a significantly negative correlation was found among students' English proficiency determined by the online simulated TOEIC and their counts of unfamiliar vocabulary (p = .001, r = -.306) and sentences (p = .000, r = -.321). Students in both groups showed significantly better performance in all the writing parameters of the post-writing than in the pre-writing, with effect sizes ranging from 0.65 to 2.19 for Task 1 and from 0.56 to 1.65 for Task 2, such as performing a significantly higher writing score, writing more enriched content, using more high-order words and having a significantly lower probability of writing errors in their post-writing, as shown in Table 2.

Students' FRE scores of the post-writing were improved from 51.1 and 65.4 in the pre-writing to 43.1 and 61.4 in the post-writing for Tasks 1 and 2, respectively. In addition, the means of word count students borrowed from the GenAI-generated templates for Tasks 1 and 2 were 26.9 and 34.4, accounting for 25.3% and 28.6% of the mean words in their post-writing, respectively.

Table 2
Students' writing performance in the pre- and post-writing

Task & Group (*N* = no. of students)		Task 1: Letter of inquiry INQ Group (*N* = 59)		Task 2: Letter of collection COL Group (*N* = 58)	
Items		GenAI template	Mean (SD)	GenAI template	Mean (SD)
Count of unfamiliar vocabulary that students highlighted			5.4 (3.51)		5.6 (4.80)
Count of unfamiliar sentence that students highlighted			1.3 (1.22)		0.9 (1.14)
Grammarly score	Pre-writing	90	81.4 (15.7)	85	80.0 (13.1)
	Post-writing		89.7 (10.0)**		86.2 (8.9)**
	Cohen's *d*		.65		.56
Total words	Pre-writing	128	68.0 (40.2)	158	71.9 (26.7)
	Post-writing		106.2 (49.2)**		120.1 (35.2)**
	Cohen's *d*		.93		1.56
K1 words	Pre-writing	95	51.3 (31.2)	127	62.3 (23.1)
	Post-writing		89.3 (36.9)**		99.8 (28.6)**
	Cohen's *d*		.85		1.45
K2 words	Pre-writing	8	4.7 (2.5)	11	3.2 (2.3)
	Post-writing		7.2 (3.3)**		7.5 (3.5)**
	Cohen's *d*		.87		1.44
AWL+ off-list words	Pre-writing	25	12.0 (9.7)	20	6.3 (3.2)
	Post-writing		18.7 (11.0)**		13.7 (7.5)**
	Cohen's *d*		.66		1.38
Different words	Pre-writing	94	49.3 (21.2)	102	52.2 (15.6)
	Post-writing		74.5 (25.7)**		80.2 (18.2)**
	Cohen's *d*		2.19		1.65
Lexical words	Pre-writing	69	37.6 (21.4)	76	35.1 (14.3)
	Post-writing		57.8 (27.4)**		58.2 (17.2)**
	Cohen's *d*		.83		1.47
FRE score	Pre-writing	36.8	51.1 (12.2)	60.5	65.4 (10.2)
	Post-writing		43.1 (11.2)**		61.5 (7.7)**
	Cohen's *d*		.68		.44
Possibility of writing errors per words	Pre-writing	0.0000	.0556 (.0527)	0.0000	.0569 (.0549)
	Post-writing		.0186 (.0265)**		.0195 (.0266)**
	Cohen's *d*		.93		.92
Mean of word count increased in the post-writing (percentage)			38.2 (36.0%)		48.2 (40.1%)
Mean of word count borrowed from the GenAI template in the post-writing (percentage)			26.9 (25.3%)		34.4 (28.6%)

$^*p < .05$. $^{**}p < .01$.

The count and percentage of the professional points mentioned in the students' pre- and post-writing for the two tasks are shown in Table 3:

- Task 1: In addition to 11 professional points (IPP 1–IPP 11) mentioned in the GenAI template, two additional professional points, asking for after-sales service (IPP 12) and a discount (IPP 13), were mentioned by few students, respectively. A significant difference was found in the mean counts of the professional points between students' pre- (*M* = 3.7, 28.3%) and post-writing (*M* = 7.0, 54.1%), with a large size effect of 1.52. Except for IPP 2 almost mentioned in students' pre- and post-writing, they showed significantly higher counts in the other 10 professional points in their post-writing than their pre-writing, with effect sizes ranging from 0.31 to 1.04.

- Task 2: In addition to 8 professional points (CPP 1–CPP 8) mentioned in the GenAI template, two additional professional points, identifying possible problems (CPP 9) and taking legal action (CPP 10), were respectively provided in two of the three samples, and one additional professional note, expressing gratitude (CPP 11), was mentioned by few students. There was a significant difference in the mean counts of the professional points between students' pre-writing (*M* = 4.9, 44.2%) and post-writing (*M* = 6.8, 61.6%), with a large size effect of 1.39. Except for IPP 2 and CPP 3, students performed significantly higher counts in the other 6 professional points of the post-writing than in the pre-writing, with effect sizes ranging from 0.39 to 0.98.

Table 3
Count and frequency of the professional points mentioned in the students' pre- and post-writing

	Task 1: INQ Group (*N* = 59)		Task 2: COL Group (*N* = 58)	
	Professional note	Count (%)	Professional note	Count (%)
Pre-writing Post-writing Cohen's *d*	IPP 1 Brief opening	45 (76.3%) 53 (88.1%)** 0.31	CPP 1 Brief opening	53 (91.4%) 57 (98.3%)* 0.54
Pre-writing Post-writing Cohen's *d*	IPP 2 Product to be inquired	57 (96.6%) 59 (100%)	CPP 2 Reminder for not receiving the payment	55 (94.8%) 55 (94.8%)
Pre-writing Post-writing Cohen's *d*	IPP 3 Product catalogues	6 (10.2%) 24 (40.7%)** 0.76	CPP 3 No reply to the first letter of collection	39 (67.2%) 45 (77.6%)
Pre-writing Post-writing Cohen's *d*	IPP 4 Product specifications	21 (35.6%) 44 (74.6%)** 0.85	CPP 4 Overdue payment	3 (5.2%) 17 (29.3%)** 0.71
Pre-writing Post-writing Cohen's *d*	IPP 5 Product price	21 (35.6%) 37 (62.7%)** 0.56	CPP 5 Asking for payment	38 (65.5%) 51 (87.9%)** 0.49
Pre-writing Post-writing Cohen's *d*	IPP 6 Minimum order quantity	5 (8.5%) 28 (47.5%)** 1.01	CPP 6 Apology if the payment has been made	6 (10.3%) 29 (50.0%)** 0.98
Pre-writing Post-writing Cohen's *d*	IPP 7 Shipment	5 (8.5%) 28 (47.5%)** 0.99	CPP 7 Asking for a quick reply	12 (20.7%) 33 (56.9%)* 0.79
Pre-writing Post-writing Cohen's *d*	IPP 8 Order procedure	10 (16.9%) 25 (42.4%)** 0.58	CPP 8 Salutation	29 (50.0%) 46 (79.3%)** 0.69
Pre-writing Post-writing Cohen's *d*	IPP 9 Asking for a reply	14 (23.7%) 37 (62.7%)** 0.85	CPP 9 Identifying possible problems	30 (51.7%) 19 (32.8%)** 0.39
Pre-writing Post-writing Cohen's *d*	IPP 10 Expressing gratitude	20 (33.9%) 47 (79.7%)** 1.04	CPP 10 Taking legal action	16 (27.6%) 12 (20.7%)
Pre-writing Post-writing Cohen's *d*	IPP 11 Looking for cooperation opportunities	6 (10.2%) 25 (42.4%)** 0.80	CPP 11 Expressing gratitude	1 (1.7%) 3 (5.2%)
Pre-writing Post-writing Cohen's *d*	IPP 12 Asking for after-sales service	4 (6.8%) 5 (8.5%) 1.04		
Pre-writing	IPP 13	4 (6.8%)		

	Task 1: INQ Group (N = 59)		Task 2: COL Group (N = 58)	
	Professional note	Count (%)	Professional note	Count (%)
Post-writing	Asking for a	3 (5.1%)		
Cohen's d	discount	0.80		
Pre-writing	Mean of	3.7	Mean of	4.9 (44.2%)
Post-writing	professional	7.0 (54.1%)	professional	6.8 (61.6%)
Cohen's d	points mentioned	1.52	points mentioned	1.39
	by students		by students	

$^*p$ <. 05. $^{**}p$ <. 01.

RQ3: Students' perceptions

The Cronbach's alpha values and the KMO values of the questionnaire were greater than 0.7, listed in Table 4. The students' means of the first (SQ1 to SQ11) and second (TQ1 to TQ11) parts in the questionnaire were 3.93 and 3.77, respectively. The results from the independent sample t test showed that there was no significant difference in all the items of the questionnaire between students in the INQ and COL groups.

In the first part of the questionnaire, the top three highest items that students were most satisfied with were in SQ5 (providing ideas, M = 4.22), SQ8 (convenience, M = 4.19) and SQ6 (speed, M = 4.12). Students gave high means for SQ2 (vocabulary, M = 3.95), SQ3 (grammar, M = 3.85), SQ4 (sentence patterns, M = 3.91) and SQ7 (learning by referring to the GenAI-generated templates, M = 3.81). They were willing to continue to use (SQ10, M = 3.86) and recommend the use of GenAI in English writing (SQ11, M = 3.90). They gave relatively low means for SQ12 (self-critical thinking skills, M = 3.10), SQ15 (modifying the script by referring to the GenAI templates is an act of plagiarism, M = 3.08) and SQ13 (creative ability, M = 2.98). The lowest mean was in SQ14 (using GenAI against the nature of English learning, M = 2.39). The results from the Pearson correlation analysis indicated that students' means had a significantly negative correlation with the number of their unfamiliar vocabulary (r = -.214, p = .020) and sentences (r = -.343, p = .000).

The top three highest items in the second part of the questionnaire that students were most satisfied with were TQ11 (worthy of reference, M = 4.06), TQ6 (enriched content, M = 3.93), TQ4 (well-structured content, M = 3.90) and TQ7 (appropriate content, M = 3.90). The lowest mean was in TQ10 (M = 3.15, lengthy content).

Table 4
Student perceptions of the study

Part 1. Students' perceptions about the GenAI-GETF approach			
	ALL (N = 117) α = .865 KMO = .831	INQ (N = 59) α = .870 KMO = .773	COL (N = 58) α = .861 KMO = .810
Items	Mean (SD)	Mean (SD)	Mean (SD)
SQ1: I think the GenAI template is helpful in terms of "content" for English business writing.	4.01 (.749)	4.02 (.707)	4.00 (.795)
SQ2: I think the GenAI template is helpful in terms of "vocabulary" for English business writing.	3.95 (.775)	3.92 (.726)	3.98 (.827)
SQ3: I think the GenAI template is helpful in terms of "grammar" for English business writing.	3.85 (.757)	3.81 (.754)	3.90 (.756)
SQ4: I think the GenAI template is helpful in terms of "sentence patterns" in English business writing.	3.91 (.754)	3.88 (.768)	3.93 (.746)
SQ5: I think the GenAI template are helpful in "providing ideas" for English business writing.	4.22 (.756)	4.20 (.714)	4.24 (.802)
SQ6: I think modifying the script by referring to the GenAI template will be helpful in terms of "speed".	4.12 (.822)	4.05 (.818)	4.19 (.826)

SQ7: I think I have the ability of modifying the GenAI template.	3.43 (.977)	3.37 (1.015)	3.48 (.941)
SQ8: I think it is convenient to use GenAI for English business writing.	4.19 (.742)	4.17 (.746)	4.21 (.744)
SQ9: I can learn by referring to the GenAI-generated template.	3.81 (.919)	3.83 (.874)	3.79 (.969)
SQ10: I will continue to use GenAI in English writing.	3.86 (.840)	3.90 (.874)	3.83 (.881)
SQ11: I recommend other students or friends to use GenAI in English writing.	3.90 (.904)	3.88 (.873)	3.91 (.942)
Overall mean & SD of items SQ1–SQ11	3.93 (.818)	3.91 (.806)	3.95 (.839)
SQ12: I think using GenAI is not favourable for the development of self-critical thinking skills.	3.10 (1.020)	3.08 (.970)	3.12 (1.077)
SQ13: I think using GenAI is not favourable for the development of creative abilities.	2.98 (.900)	3.12 (.745)	2.84 (1.023)
SQ14: I think using GenAI creates a sense of dependence against the nature of English learning.	2.39 (1.058)	2.32 (.990)	2.47 (1.127)
SQ15: I think modifying the script by referring to the GenAI template is an act of plagiarism.	3.08 (1.060)	3.20 (1.013)	2.95 (1.099)
SQ16: I will pay attention to academic ethics and integrity when using GenAI.	3.52 (.988)	3.53 (1.023)	3.52 (.960)
Part 2. Students' opinions about the appropriateness and quality of the GenAI-generated English template			
	ALL (*N* = 117) α = .871 KMO = .888	INQ (*N* = 59) α = .881 KMO = .884	COL (*N* = 58) α = .859 KMO = .841
Items	Mean (SD)	Mean (SD)	Mean (SD)
TQ1: I think the vocabulary usage of the GenAI template is concise.	3.58 (.912)	3.54 (.877)	3.62 (.952)
TQ2: I think the grammar of the GenAI template is accurate.	3.67 (.851)	3.58 (.894)	3.76 (.802)
TQ3: I think the content of the GenAI template is easy to understand.	3.74 (.792)	3.78 (.744)	3.69 (.842)
TQ4: I think the GenAI template is well structured.	3.90 (.824)	3.93 (.807)	3.86 (.847)
TQ5: I think the professional points stressed in the GenAI template are clear.	3.84 (.798)	3.73 (.827)	3.95 (.759)
TQ6: I think the content of the GenAI template is rich.	3.93 (.828)	3.85 (.827)	4.02 (.827)
TQ7: I think the content of the GenAI template is appropriate.	3.90 (.803)	3.80 (.867)	4.00 (.725)
TQ8: I think the content of the GenAI template is clear.	3.89 (.774)	3.80 (.854)	3.98 (.737)
TQ9: I think the content of the GenAI template is coherent.	3.82 (.826)	3.80 (.805)	3.84 (.854)
TQ10: I think the content of the GenAI template is too lengthy.	3.15 (.958)	3.24 (.953)	3.05 (.963)
TQ11: I think the content of the GenAI template is worthy of reference.	4.06 (.735)	4.02 (.731)	4.10 (.742)
Overall mean & SD of TQ1–TQ11	3.77 (.827)	3.73 (.835)	3.81 (.823)

Discussion

Compared with three samples excerpted from the textbooks, the GenAI-generated English templates provided more professional points and more content with more high-order words, more lexical and different words and a higher FRE score. Students' highly positive responses to the questionnaire reinforced that the GenAI-generated templates were worthy of reference in business writing, especially in their enriched, well-structured, appropriate, and coherent content with clear professional points. Students also considered GenAI to be a convenient and useful tool for enhancing their writing performance, as mentioned in recent studies (Baskara, 2023; Song & Song, 2023; Yan, 2023). Through learning with the GenAI-GETF approach, students' linguistic and content-based knowledge in business writing (such as writing more enriched content, using more high-order words, having a lower probability of writing errors, performing a higher writing score, improving their FRE score towards the level of the GenAI-generated templates and presenting higher counts of professional points) were significantly improved, consistent with recent findings on academic writing by using GenAI (Song & Song, 2023; Teng, 2024; Tsai et al., 2024; Yan, 2023). The results from the Pearson analysis indicated that students who highlighted more unfamiliar vocabulary and sentences in the learning process had a significantly more positive satisfaction with the GenAI-GETF approach. The results reveal that making good use of technological tools in language learning was especially helpful for low-proficient students who needed more improvement on content-based and linguistic knowledge, echoing Tsai (2022, 2025). However, the percentage of some professional points was still lower than 50% in students' post-writing, which can help them understand what should be strengthened further.

Students who had a lower possibility of writing errors per word had significantly better performance in their post-writing, such as writing more content (K1: $r = .295$, $p = .001$), using more high-order words (AWL and off-list words: $r = .262$, $p = .004$), delivering more different words ($r = .352$, $p = .000$) and lexical words ($r = .307$, $p = .001$) words, performing higher scores ($r = .705$, $p = .000$) and presenting higher counts of professional points ($r = .352$, $p = .000$). These results indicate that there were significant correlations and consistent patterns among the writing parameters measured by the qualitative and quantitative assessments in this study.

As more than 93% of students had not taken the course of English business writing, the ideas provided from the GenAI-generated templates were the major reference sources for students to revise their post-writing, which could be seen in SQ5 ($M = 4.22$), the highest mean of the questionnaire. Furthermore, it may be the reason for thinking that using GenAI was not conducive to cultivating their own critical thinking ability ($M = 3.10$, SQ9) and creativity ($M = 2.98$, SQ10). A further study could be conducted in which students freely use GenAI to complete their writing task and then copy and paste their writing process into a Microsoft Word file: this would support deeper investigation on how students use and modify GenAI prompts to complete their writing as well as their critical thinking ability to solve the problems and difficulties they encounter.

Although students had a neutral attitude towards the fact that modifying GenAI templates could be an act of plagiarism, they would pay attention to academic ethics and integrity when using GenAI. This attitude was seen in the percentages of their borrowed words from the GenAI templates in Tasks 1 and 2, which were 25.3% and 28.6%, being within the range of 20%–30%, a similarity standard recognised in dissertations in universities in Taiwan (Department of Electrical Engineering, 2025; Graduate Institute of Library & Information Science, 2025; National Tsing Hua University, 2025). As only word-level comparison was checked by using the software self-developed in this study, the function involving sentence- and paragraph-level comparisons can be further developed. In addition, software like Turnitin and Plagiarism Checker can be integrated into EFL writing instruction to help students recognise the importance of respecting intellectual property rights, thereby reducing or preventing plagiarism.

According to the basic concept of process-based writing (Dokchandra, 2018; Guo, et al., 2021), the GenAI-GETF approach with a five-step learning model was successfully implemented in EFL business writing. Its implementation met the four requirements based on the research-backed principles in the SLA field:

input, output, interaction and feedback (Nielson, 2022). In addition, I proposed some solutions to common problems caused by using technology For example, while referring to the GenAI-generated templates in the editing and revising stages, students could use familiar AI-based tools for receiving immediate and personalised feedback to improve their comprehensible input and output, as mentioned in some studies in the literature (Loncar et al., 2021; Lu & Zeng, 2025). These student-centred activities could promote students' consciousness and opportunities of learning vocabulary knowledge, semantic and syntactic structure, or content knowledge and allow them to notice and comprehend unfamiliar linguistic forms of the target content. These features can help enhance meaningful content-based and linguistic interaction and reduce superficial interactions and distraction from learning tasks (Darkoah et al., 2024; Shadiev & Yang, 2020). It may be the reason why students had a tendency not to think that using GenAI could create a sense of dependency against the nature of English learning, as seen in SQ14.

Implications and limitations

Based on the findings of the study, some implications are proposed to ensure the optimal utilisation of the GenAI-GETF approach for online learning: First, according to different contexts, GenAI can immediately generate purposeful templates, including content-based and linguistic knowledge for students' reference. Second, while learning with the GenAI-generated templates, students can use available AI-based tools to monitor, check, notice, highlight, translate, identify and evaluate what should be learned, corrected, modified and reinforced. These student-centred strategies can enhance students' interaction with the target content and help them produce well-structured and meaningful output. Third, based on the curriculum objectives, EFL teachers can make good use of the features of technological tools to prepare instructional materials, design and administer linguistic activities and strategies for effectively enhancing students' interaction and use different assessments to measure their performance in a multi-faceted and objective manner. Fourth, the GenAI-GETF approach combined with the use of AI-based tools and software in this study proposed some possible solutions to the common problems encountered in the online learning process, such as the risk of inaccurate feedback, plagiarism, superficial interactions and distraction.

Some limitations of this study are as follows: First, as the results of this study were obtained by one-time achievement only, more research regarding periodic experiments should be conducted to investigate and validate the long-term effect of the GenAI-GETF approach. Second, a study with two control groups can be conducted further: one control group with templates written by EFL teachers based on real workplace cases and the other with templates excerpted from textbooks to compare with the effectiveness of the GenAI-GETF approach employed in this study. Third, the participants were non-English major EFL students from a university in Taiwan so the results might not be generalised globally. Finally, the GenAI-GETF approach can be further expanded to other English skills or subject courses with a larger group of students to investigate the possible pedagogical benefits and impact of GenAI.

Conclusion

The GenAI-GETF approach with a five-step learning model was effectively implemented in two business contexts for a total of 117 non-English major EFL students in two groups. Results indicate that students showed significantly better performance in all the writing parameters of the post-writing than the pre-writing, such as writing more enriched content, using more high-order words, having a lower probability of writing errors and performing better scores in their post-writing, with effect sizes ranging from 0.65 to 2.19 for Task 1 and from 0.56 to 1.65 for Task 2. Students also showed high satisfaction with their learning through the GenAI-GETF approach and had positive perceptions towards the appropriateness and quality of the content provided by GenAI-generated templates. These results indicated that the use of AI-based tools combined with student-centred activities assigned by the teacher in the GenAI-GETF approach provided students with multimodal input, scaffolding and feedback so that their interaction with the target content can be enhanced through computers for promoting their engagement and improving their writing performance. To further understand possible advantages and impacts with its implementation, more studies with various instructional modes, activities, strategies and tasks in different courses and domains should be conducted to investigate and evaluate the learning effectiveness and perceptions of students with different educational backgrounds.

Acknowledgements

This work was partially supported by National Science and Technology Council, Taiwan, R.O.C. (grant number NSTC 114-2410-H-992-018).

References

Allen, T., & Mizumoto, A. (2024). ChatGPT over my friends: Japanese EFL learners' preferences for editing and proofreading strategies. *RELC Journal*. https://doi.org/10.1177/00336882241262533

Arifin, S. (2017). L2 writing strategies used by EFL graduate students. *Journal of ELT Research, 2*(2), 115–129. https://doi.org/10.22236/JER_Vol2Issue2pp115-129

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

Baskara, F. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5*(1), 44–55. https://doi.org/10.31763/ijele.v5i1.858

Chapelle, C. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. Language Learning & Technology, 2(1), 21–39. https://doi.org/10125/25030

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Darkoah, C., Abunya, L. N., Owusu, E., & Faleke, V. O. (2024). Investigating the effectiveness of technology-enhanced language learning tools in supporting second language acquisition. *Journal of Education and Learning Technology, 5*(10), 427–439. https://doi.org/10.38159/jelt.20245104

Department of Electrical Engineering. (2025). 論文比對規定 [Thesis comparison regulations]. National Taiwan University of Science and Technology. https://www.ee.ntust.edu.tw/p/412-1027-11036.php?Lang=zh-tw

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies, 29*, 14151–14203. https://doi.org/10.1007/s10639-023-12402-3

Dizon, G., & Gayed, J. M. (2024). A systematic review of Grammarly in L2 English writing contexts. *Cogent Education, 11*(1). https://doi.org/10.1080/2331186X.2024.2397882

Dokchandra, D. (2018). The effects of process writing approach on performance of an overcrowded EFL writing class at a university in Thailand. *KnE Social Sciences*, *3*(4), 191–206. https://doi.org/10.18502/kss.v3i4.1931

Druce, P. M. (2012). Attitude to the use of L1 and translation in second language teaching and learning. *Journal of Second Language Teaching and Research, 2*(1), 60–86. http://pops.uclan.ac.uk/index.php/jsltr/article/view/82/30

El Harrath, A. (2024). The intersection of computer-assisted-language-learning and SLA. *International Journal of Linguistics, Literature and Translation, 7*(6), 15–19. https://doi.org/10.32996/ijllt.2024.7.6.3

Graduate Institute of Library & Information Science. (2024). 學位論文品質與管考準則 [Regulations on thesis quality and management]. National Chung Hsing University. https://www.gilis.nchu.edu.tw/upload/regulations/25011314095400000002.pdf

Guo, W., Bai, B., & Song, H. (2021). Influences of process-based instruction on students' use of self-regulated learning strategies in EFL writing. *System, 101*, Article 102578. https://doi.org/10.1016/j.system.2021.102578

Harunasari, S. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technologies, 39*(2), 357–368. https://doi.org/10.52155/ijpsat.v39.2.5516

Holt, R., & Sampson, N. (2004). *International business correspondence for all occasions*. Macmillan Publishers.

Isbell, D. R., Dixon, D. H., & Plonsky, L. (2022). Research methods at the intersection of technology and SLA: Advances and challenges. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of second language acquisition and technology* (pp. 65–79). Routledge. https://doi.org/10.4324/9781351117586-7

Isma, A., Putri, A. M. J., & Sardi, A. (2024). Examining university students' business English writing performance: Frequent errors and pitfalls. *Research and Innovation in Applied Linguistics, 2*(2), 79–94. https://doi.org/10.31963/rial.v2i2.4446

Kayalı, B., Yavuz, M., Balat, Ş., & Çalışan, M. (2023). Investigation of student experiences with ChatGPT-supported online learning applications in higher education. Australasian Journal of Educational Technology, 39(5), 20–39. https://doi.org/10.14742/ajet.8915

Li, L., & Kim, M. (2024). It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*, *40*(1), 1–18. https://doi.org/10.14742/ajet.8821

Liang, L. (2025). The impact of using DeepL translator on Chinese EFL students' story writing. *Journal of China Computer-Assisted Language Learning, 5*(1),56–93. https://doi.org/10.1515/jccall-2024-0009

Loncar, M., Schams, W., & Liang, J.-S. (2021). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015-2019. Computer Assisted Language Learning, 36(4), 722–784. https://doi.org/10.1080/09588221.2021.1943452

Lu. D., & Zeng, Y. (2025). Exploring the use of ChatGPT-generated model texts as a feedback instrument: EFL students' text quality and perceptions. Innovation in Language Learning and Teaching, 1–21. https://doi.org/10.1080/17501229.2025.2525341

Macintosh, D. (1994). English for business (6th ed.). Book Marketing Ltd.

Mai, L. (2025). Using genre-based approach to teach business email writing in a Vietnamese university: Making complaints. *Vietnam Journal of Education*, *9*(1), 46–60 https://doi.org/10.52296/vje.2025.554

McLeod, S. A. (2023, July 31). *What does effect size tell you?* Simply psychology. Retrieved December 15, 2025, from https://www.simplypsychology.org/effect-size.html

National Tsing Hua University. (2025). 博學位考試需備論文相似度比對報告 [Thesis similarity comparison report]. https://mba.site.nthu.edu.tw/p/412-1370-17732.php

Nguyen, T. T. H. (2023). EFL teachers' perspectives toward the use of ChatGPT in writing classes: A case study at Van Lang university. *International Journal of Language Instruction*, *2*(3), 1–47. https://doi.org/10.54855/ijli.23231

Nielson, K. B. (2022). Language learning in online environments. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of second language acquisition and technology* (pp. 233–244). Routledge. https://doi.org/10.4324/9781351117586-7

Pham, T., Nguyen, T. B., Ha, S., & Nguyen Ngoc, N. T. (2023). Digital transformation in engineering education: Exploring the potential of AI-assisted learning. *Australasian Journal of Educational Technology*, *39*(5), 1–19. https://doi.org/10.14742/ajet.8825

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, *12*, Article 524. https://doi.org/10.3390/su12020524

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Tahriri, A., Shabani, M., & Zokaei, S. (2016). EFL learners' attitudes toward writing instruction based on critical language awareness. *Theory and Practice in Language Studies*, *6*(1), 127–133. https://doi.org/10.17507/tpls.0601.17

Taylor, S. (2012). *Model business letters, emails & other business documents* (7th ed.). Financial Times Press.

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, *6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Tsai, C. Y., Lin, Y. T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, *29*, 22427–22445. https://doi.org/10.1007/s10639-024-12722-y

Tsai, S. C. (2017). Effectiveness of ESL students' performance by computational assessment and role of reading strategies in courseware-implemented business translation tasks. *Computer Assisted Language Learning*, *30*(6), 474–487. https://doi.org/10.1080/09588221.2017.1313744

Tsai, S. C. (2019). Implementing interactive courseware into EFL business writing: computational assessment and learning satisfaction. *Interactive Learning Environments*, *27*(1), 46–61, https://doi.org/10.1080/10494820.2018.1451896

Tsai, S. C. (2022). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer Assisted Language Learning*, *35*(5-6), 1250–1272. https://doi.org/10.1080/09588221.2020.1799412

Tsai, S.-C. (2025). Implementing the digital multimodality of a Google Translate-assisted approach to online academic EFL reading. *Australasian Journal of Educational Technology*, *41*(2), 89–104. https://doi.org/10.14742/ajet.10120

Weijen, D., Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing, 18*(4), 235–250. https://doi.org/10.1016/j.jslw.2009.06.003

Wilkinson, C., Oppert, M., & Owen, M. (2024). Investigating academics' attitudes towards ChatGPT: A qualitative study. *Australasian Journal of Educational Technology*, *40*(4), 105–120. https://doi.org/10.14742/ajet.9456

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

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*, 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Corresponding author: Shu-Chiao Tsai, achiao@nkust.edu.tw

Please cite as: Tsai, S.-C. (2025). Online EFL business writing with GenAI-generated templates: Students' performance and perceptions. *Australasian Journal of Educational Technology*, *41*(6), 82–97. https://doi.org/10.14742/ajet.10743