



## **Peer-assisted vs. AI-assisted Collaborative Writing: Production Quality and Learners' Perceptions**

**Sahar Zahed-Alavi**

*Language and Linguistics Department, University of Bojnord, Bojnord, Iran.*

[s.zahedalavi@ub.ac.ir](mailto:s.zahedalavi@ub.ac.ir)

### **ARTICLE INFO:**

**Received date:**

**2025.08.31**

**Accepted date:**

**2025.10.20**

**Print ISSN:** 2251-7995

**Online ISSN:** 2676-6876

### **Keywords:**

AI-Assisted Writing, Peer-Assisted Writing, Perceptions



### **Abstract**

Although peer-assisted collaborative writing has been examined extensively, the development of artificial intelligence (AI) highlights the importance of investigating how AI might affect the way students write and perceive writing. Drawing on the sociocultural theory, the present mixed-methods study examined learners' integration of AI in their writing process. The participants included 40 TEFL students, randomly assigned to either the peer-assisted collaborative writing or AI-assisted collaborative writing groups. In the first phase of the study, a pretest and a posttest were administered to both groups. They were analyzed in terms of different measures of complexity, accuracy, and fluency features (CAF). A series of paired and independent t-tests was used to compare the two groups in each measure. Although there was no difference between the groups in terms of complexity and fluency, learners in the peer-assisted group outperformed those in the AI-assisted group in terms of accuracy measures. In the second phase of the study, the participants in both groups filled in an open-ended questionnaire on their attitudes toward their writing experiences. It was found that learners in the peer-scaffolded group were more satisfied with their writing and experienced less boredom during writing.

**Citation:** Zahed-Alavi, S. (2025). Peer-assisted vs. AI-assisted Collaborative Writing: Production Quality and Learners' Perceptions. *Journal of English Language Teaching and Learning*, 17 (36), 341-362.  
DOI:10.22034/elt.2025.68919.2818

## Introduction

The application of Artificial Intelligence (AI) and digital technologies in education and language learning has been globally expanded in recent years (Wang, 2024). AI is essential for enhancing the efficiency, personalization, and engagement of education and language acquisition (Escalante, 2023). The implementation of AI and technologies has resulted in a progressively responsive and adaptive learning environment, advancing towards an ideal in which AI functions as a facilitator for improved learning outcomes (Kim et al., 2022). Although technological improvements have transformed education, AI plays a particularly essential role in redefining collaborative writing and scaffolding in learning contexts (Özçelik & Yangn Ekş, 2024).

Collaborative writing encompasses not only the final text as the outcome, but also involves shared cognition, including the generation of ideas, writing, concern for grammatical points, and editing (Anshu & Yesuf, 2022). Literature has indicated the possible advantages of collaborative writing such as providing students with more learning opportunities (Pham, 2021), co-constructing knowledge (Qiu & Lee, 2020), discussing language elements (Storch, 2013), being aware of the gaps in their writing knowledge through interaction (Bueno-Alastuey et al., 2022), and getting rid of anxiety or stress of writing (Anshu & Yesuf, 2022). In addition, research on AI-based scaffolding is rapidly advancing, highlighting its growing significance in providing support for various aspects of the writing process, including feedback, linguistic assistance, and idea generation (Wang, 2024).

However, despite these advancements, the comparative responsibilities of peer scaffolding and AI scaffolding remain inadequately comprehended and insufficiently analyzed (Escalante et al., 2023; Nguyen et al., 2024). Indeed, despite the significant developments and growing role of AI-based assistance in providing feedback, language support, and idea generation, there is still uncertainty about which type of scaffolding (i.e., peer or AI-based) improves writing quality, learner engagement, or satisfaction, and these two forms of scaffolding are not adequately compared (Li, 2023). Our comprehension of the most effective ways to scaffold collaborative writing is severely hampered by this ambiguity (Escalante et al., 2023). Misguided dependence on either peers or AI without evidence could jeopardize students' autonomy, writing quality, or classroom effectiveness. There is an urgent need for thorough research to elucidate the relative benefits and limitations of writing supported by peer and AI.

To fill the mentioned gap, using a mixed-methods design, this study will compare peer-assisted collaborative writing vs. AI-assisted collaborative writing in terms of the quality of the learners' writings and their perceptions of their writing experiences. Thus, the following research inquiries are raised.

1. What are the differences between peer-scaffolded collaborative writing and AI-scaffolded collaborative writing groups in terms of writing proficiency features of complexity, accuracy, and fluency (CAF)?

2. How do the EFL learners in peer-scaffolded collaborative writing and AI-scaffolded collaborative writing groups perceive their writing experiences?

## Review of Literature

### Peer-assisted collaborative writing

In the literature, collaborative writing has been examined from three different perspectives. The first perspective has investigated the influence of collaborative writing on the characteristics of learners produced texts. Interestingly, studies in this perspective have yielded different results due to variations in certain aspects. Although a group of studies considered group writing (Mei et al., 2024; Moxie et al., 2025; Peng et al., 2022), others focused on dyadic writing activities (Dobao, 2013; Hiromori, 2021; Zhang, 2022). In addition, different studies have examined collaborative writing in different phases; in other words, some studies have examined collaborative writing just in the prewriting phases (Hsu, 2025; Pham, 2021) while others have considered it throughout the whole writing process (Bueno-Alastuey et al., 2022; Chen, 2019). Furthermore, different studies analyzed the quality of the texts through various data analysis procedures. For instance, Pham (2021) found that learners who involved in collaborative tasks improved in terms of fluency; Bueno-Alastuey et al. (2022) took the three criteria of complexity, accuracy, and fluency (CAF) into account and found that learners in collaborative groups improved in terms of accuracy. Sang and Zou (2023) analyzed the learners' texts in terms of two measures of complexity and accuracy and found that learners who collaborated improved in terms of accuracy. Furthermore, Chen (2019) studied the quality of students' productions in terms of content, grammar, organization, mechanics, and vocabulary. He found that students who collaborated in writing outperformed in terms of grammar, vocabulary, and organization.

The second perspective has focused on the learners' interactions while participating in collaborative writing tasks. The interactions were analyzed in terms of the patterns (Su et al., 2024; Styati, 2018), the type of language-related episodes (Zhang, 2019), and the nature of peers' scaffolding (Zhang, 2019).

Finally, the third perspective has focused on the way learners had perceived collaborative writing tasks. They mostly agreed that learners enjoyed the activities and believed that they improved their writing skill and self-confidence. They found that although learners faced many challenges in collaborative activities, collaborative activities enhanced their teamwork skills and provided them with self-reflective skills (Sridharan, 2024), offered affordance opportunities (Alqasham, 2022), helped them to present their knowledge and opinions and be more active, creative, and accurate grammatically and lexically (Dobao & Blum, 2014), motivated them to write, saved their time, resulted in higher quality productions, assisted them in being aware of the time limit in which they had to perform, helped them to share their experiences (Ajmi & Ali, 2014), and improved their self-esteem and their writing skill. It also helped them in generating ideas, discussing, planning, drafting, and providing immediate feedback (Shehadeh, 2011). However, only a few students in online collaborative writing context had negative perceptions of collaborative writing tasks; it could be attributed to their

negative emotions towards the online platform of collaborative tasks which demotivated them from writing (Alqasham, 2022).

### **AI-assisted collaborative writing**

AI-assisted collaborative writing illustrates a shift from individual activity to social process distributed across individuals and tools (Wang, 2024). AI tools are increasingly considered as partners, mentors, and teachers in collaborative endeavor (Song & Song, 2023). Studies suggest that AI assistance can help learners improve the quality of their writing in different ways (Gayed et al., 2022). First, it can improve paragraph organization, cohesion, coherence, and grammatical structure of the texts (Kim et al., 2022). Second, it can reduce the mental load of learners and help them focus on higher-order areas such as content and argument development (Song & Song, 2023). Thus, it can reduce learners' anxiety level (Darvishi et al., 2024). Third, it assists learners in brainstorming and idea generation to start writing (Levine et al., 2024). Fourth, it offers learners instant feedback (Wang, 2024). However, it is found that excessive use of AI may lose the creativity and critical thinking of learners, leading to expectations of excessive external support rather than relying on their capabilities (Escalante, et al., 2023; Malik et al., 2023).

Despite growing research on peer-assisted and AI-assisted writing, few studies have directly compared the two approaches in terms of writing quality and learners' perceptions (Kim et al., 2022; Nguyen et al., 2024). Moreover, prior research has often treated AI as a passive tool correcting learners' structural error (Wang, 2024). The present study addresses these gaps by investigating how EFL learners engage in peer-assisted versus AI-assisted collaborative writing, using structured AI prompts to support co-construction and negotiation. By comparing both textual outcomes and learner perceptions, this study aims to provide empirical evidence on the pedagogical affordances and limitations of AI as a collaborative partner, offering practical and theoretical insights for EFL writing instruction.

## **Methods**

### **Participants**

The participants included 40 EFL students at a university in Iran. They were in the age range of 19-23 years old. They were at the intermediate proficiency level based on the results of the Oxford Placement Test (2001). During their two-credit course on Academic Writing in their third semester, the instructor (a PhD assistant professor with 15 years of experience teaching EFL students) taught the qualities of well-formed paragraphs and introduced them to different types of paragraphs. Although both groups received the same instructions on writing from the same instructor, the instructor assigned the students to either peer-assisted collaborative writing or AI-assisted collaborative writing groups randomly. Thus, each group included 20 students.

### **Instruments**

#### ***Pretests and Posttests***

The materials included 80 writings. They included 40 writings written individually by the

participants in the peer-assisted and AI-assisted groups in the pretest phase and 40 writings written individually by the participants in the peer-assisted and AI-assisted groups in the posttest phase. The prompts were the same for both groups.

### ***Questionnaire on Learners' Perceptions***

The participants in both peer-assisted and AI-assisted groups filled in an open-ended questionnaire on their attitudes toward their writing experiences. This questionnaire was adopted from Shehadeh (2011). The objective of the survey was to detect the learners' opinions on different aspects of their writing experiences. Particularly, the researcher tried to see how the learners found their writing experiences, what they liked about the experience, what the hardest part of the experience was, what their perception of writing was prior to the activity, if the activity affected their confidence in writing, if the activity affected their language skills other than writing, and finally, if the learners were willing to experience similar activities in the future.

### ***Reflective Log***

A reflective log was used to capture learners' engagement during the writing tasks across both the AI-assisted and peer-assisted groups. After performing each task, all participants completed a short reflection describing (a) the stages of writing in which they used assistance (brainstorming, drafting, revising/editing), (b) which suggestions or feedback they accepted or rejected and why, and (c) how the assistance influenced their writing decisions. These logs were just designed to ensure learners' active involvement in the collaborative writing process and to verify treatment fidelity.

### ***Data Collection***

At the beginning of the investigation, the participants in both groups were asked to write on a task individually (i.e., pretests). Then, both groups received similar instructions on how to write academically; they were taught different components of a well-written paragraph (i.e., topic sentence, body, and conclusion), the quality of a well-written paragraph (i.e., unity, cohesion, and coherence), and different types of paragraphs. Both groups were also informed of what collaborative writing is and the way it is conducted through all stages of writing (i.e., planning, drafting, revising, and editing). The participants in the peer-assisted writing group did 10 writing tasks in collaboration with a partner during the semester, and the participants in the AI-assisted group did 10 writing tasks using ChatGPT. The participants in the AI-assisted group were instructed to use AI collaboratively by a pre-task training session on structured prompting using four key elements (i.e., role, context, task, and output). This framework guided students to treat the AI as an interactive writing partner collaborating in all stages of writing (i.e., planning, drafting, revising, and editing) rather than a passive correction tool (see Appendix A). In addition, to ensure learners' active engagement in the collaborative writing process, two procedures were implemented. First, the instructor monitored their performance during the tasks. Second, after performing each task, learners completed a short reflection describing the stages of writing in which they used assistance (brainstorming, drafting, revising/editing) and which suggestions or feedback they accepted or rejected and why.

At the end of the semester, the participants in both groups were asked to write on a task individually (i.e., posttests). The time gap between the pretest and posttest was 14 weeks. Then, the questionnaire on learners' attitudes toward their writing experiences was distributed among the learners in both groups at the end of the semester. The survey was administered in the learners' classroom by their instructor.

## Data Analysis

The qualitative analysis of the reflective logs written by learners during the semester was primarily meaning-oriented in all writing phases in the AI-assisted group, thereby ensuring treatment fidelity. In addition, learners in the peer-assisted group collaborated in all writing phases.

The participants' writings in the pretest and posttest phases were examined through CAF measures. To provide a comprehensive picture of writing proficiency, as suggested by Norris and Ortega (2009), multiple measures were examined for each of CAF features; complexity was investigated through examining the ratio of clauses to T-units (Foster & Skehan, 1998) and the ratio of dependent clauses to all clauses (Wolfe-Quintero et al., 1998), accuracy was investigated through considering the ratio of error-free T-units to all T-units and the ratio of error-free clauses to all clauses (Dobao, 2013), and fluency was examined through number of words, T-units, and clauses in the text (Wigglesworth & Storch, 2009).

Intra-rater and inter-rater reliability were examined. After five weeks, the researcher estimated the units needed to analyze CAF measures and conducted Kappa agreement coefficients. They turned out to be .93 for the number of T-units, .92 for the number of dependent clauses, .91 for clauses, .92 for error-free clauses, and .94 for error-free T-units. In addition, the researcher asked a colleague who holds a PhD in TEFL to examine the mentioned units in 20 writings. The agreement indices turned out to be .93, .91, .90, .89, and .90, respectively.

Then, independent-samples t-tests comparing the groups in terms of the subcategories of CAF measures in the pretest were run to investigate the homogeneity of the groups. Consequently, paired-samples t-tests comparing the performance of the peer-assisted group in the pretest and the posttest in terms of the subcategories of CAF measures and paired-samples t-tests comparing the performances of the AI-assisted group in the pretest and the posttest in terms of the subcategories of CAF measures were conducted to examine the improvement of the participants in each group during the semester. Finally, independent-samples t-tests comparing both groups in terms of the subcategories of CAF measures in the posttest were run to explore the final performances of the two groups. To control Type 1 errors, the Bonferroni adjustment was estimated by dividing the alpha level of .05 by the number of comparisons for each measure. As a result, the alpha level turned out to be .016 for interpreting the fluency measures, and .025 for interpreting the complexity and accuracy measures.

Moreover, the analysis of the qualitative data in terms of the learners' attitudes toward their writing experiences involved open coding, axial coding, and selective coding following Ary et al. (2010). In other words, the transcribed data were read, and the concepts presented by the

learners were extracted through open coding. In the axial coding, the concepts were further subsumed under a limited number of categories based on their relationships with each other and the category under which they were placed. In the selective phase, the categories which emerged were further subsumed under some broader categories (i.e., themes). Finally, common themes shared in the learners' views were tabulated.

Concerning the reliability of coding, to estimate the intra-coder reliability, the researcher randomly selected 10 questionnaires from among those of each group (i.e., 10 questionnaires were selected from the peer-assisted group's data, and 10 questionnaires were selected from the AI-assisted group's data). The researcher investigated them again after five weeks. Using the Kappa agreement index, the intra-coder reliability was found to be .96.

The selected questionnaires were coded again by another colleague experienced in this area (with a PhD degree in TEFL) to estimate the inter-coder reliability. The Kappa coefficient for the inter-coder reliability was found to be .88. It is worth mentioning that the disagreements were resolved through discussions.

## Results and Discussion

### Learners' Performance in Terms of CAF Measures

Table 1 presents the descriptive and inferential statistics examining the homogeneity of the learners in the peer-assisted and AI-assisted groups in terms of three measures of CAF in the pretest. Based on the Bonferroni adjustment to control Type 1 errors (i.e, the alpha levels of .016 for interpreting the fluency measures, and .025 for interpreting the complexity and accuracy measures), there were not any significant differences between the groups in terms of the fluency measures as far as the average number of words per text ( $t = -2.5$ ,  $df = 38$ ,  $p > .016$ ), the number of T-units per text ( $t = -1.21$ ,  $df = 38$ ,  $p > .016$ ), and the number of clauses per text ( $t = -.264$ ,  $df = 38$ ,  $p > .016$ ) are concerned. Concerning the complexity measures, there were not any significant differences between the two groups in terms of clauses per T-units ( $t = 1.13$ ,  $df = 38$ ,  $p > .025$ ) and dependent clauses percentage ( $t = 1.15$ ,  $df = 38$ ,  $p > .025$ ). Concerning the accuracy feature, there were no significant differences between the two groups in terms of the percentage of error-free T-units ( $t = -1.11$ ,  $df = 38$ ,  $p > .025$ ) and error-free-clauses percentage ( $t = -.682$ ,  $df = 38$ ,  $p > .025$ ). Therefore, learners in both groups were homogeneous in terms of the subcategories of CAF features at the outset.

**Table 1.** *Independent-Samples T-Tests Comparing the Peer-Assisted and AI-Assisted Groups in Terms of the Subcategories of CAF Measures in the Pretest*

	Levene's test for equality of variances				t-test		
	Mean	SD.	F.	Sig.	t	df	Sig.
<b>Fluency measures</b>							
<i>Words</i>							
AI-assisted Experiments	132.55	33.88	.220	.642	-2.30	38	.027
	158.15	36.45					
<i>T-units</i>							
AI-assisted Experiments	10	2.20	.615	.438	-1.21	38	.232
	10.95	2.72					
<i>Clauses</i>							
AI-assisted Experiments	16.1	4.64	2.66	.111	-.264	38	.793
	16.45	3.67					
<b>Complexity measures</b>							
<i>Clauses per T-unit</i>							
AI-assisted Experiments	1.60	.31	1.02	.318	1.13	38	.265
	1.50	.24					
<i>Dependent clauses percentage</i>							
AI-assisted Experiments	37.55	11.28	.152	.699	1.15	38	.257
	33.57	10.71					
<b>Accuracy measures</b>							
<i>Error-free T-units percentage</i>							
AI-assisted Experiments	54.50	16.15	.690	.411	-1.11	38	.274
	60.57	18.39					
<i>Error-free clauses percentage</i>							
AI-assisted Experiments	68.26	12.93	.121	.730	-.682	38	.499
	71.25	14.74					

Table 2 presents the results of paired-sample t-tests used to compare the performances of the AI-assisted group in the pretest and posttest in terms of the subcategories of CAF measures. Based on the Bonferroni adjustment, there were significant differences between the two sets of scores at the level of .016 in the number of words ( $t = -3.96$ ,  $df = 19$ ,  $sig < .016$ ), the number of T-units per text ( $t = -3.41$ ,  $df = 19$ ,  $sig < .016$ ), and the number of clauses per text ( $t = -3.06$ ,  $df = 19$ ,  $sig < .016$ ). As a result, it may be concluded that there was an increase in the AI-assisted group's scores on the posttest in terms of the three fluency measures. The eta squared statistic (.47, .42, and .37 for the number of words, T-units, and clauses, respectively) indicated large effect sizes (Cohen, 1988).

However, concerning the comparison of the AI-assisted group's performance in terms of the complexity measures in the pretest and posttest, at the significance level of .025, there were not any significant differences between the two sets of scores in terms of clauses per T-unit ( $t = 1.2$ ,  $df = 19$ ,  $sig > .025$ ) and dependent clauses percentage ( $t = 1.29$ ,  $sig > .025$ ). Similarly, there were not any significant differences between the two sets of scores in terms of error-free T-units percentage ( $t = -.181$ ,  $df = 19$ ,  $sig > .025$ ) and error-free clauses percentage ( $t = -.050$ ,  $df = 19$ ,  $sig > .025$ ). Therefore, the learners in the AI-assisted group did not improve in terms of

the complexity and accuracy measures.

**Table 2.** Paired-Samples T-Tests Comparing the Performance of the AI-Assisted Group in the Pretest and the Posttest in Terms of the Subcategories of CAF Measures

	N	M	SD	T	df	Sig.
<b>Fluency measures</b>						
<i>Words</i>						
Pretest	20	132.55	33.88	-3.96	19	.001
Posttest	20	182.45	45.29			
<i>T-units</i>						
Pretest	20	10	2.20	-3.41	19	.003
Posttest	20	15.20	5.93			
<i>Clauses</i>						
Pretest	20	16.10	4.67	-3.06	19	.006
Posttest	20	22.35	7.91			
<b>Complexity measures</b>						
<i>Clauses per T-unit</i>						
Pretest	20	1.60	.31	1.2	19	.249
Posttest	20	1.49	.30			
<i>Dependent clauses percentage</i>						
Pretest	20	37.58	11.28	1.29	19	.212
Posttest	20	32.28	13.47			
<b>Accuracy measures</b>						
<i>Error-free T-units percentage</i>						
Pretest	20	54.50	16.15	-.181	19	.858
Posttest	20	55.22	16.96			
<i>Error-free clauses percentage</i>						
Pretest	20	68.26	12.93	-.050	19	.961
Posttest	20	68.37	14.05			

Moreover, the results of paired-samples t-test run to compare the performances of the peer-assisted group in the pretest and the posttest in terms of the subcategories of CAF measures are presented in Table 3. Considering the Bonferroni adjustment, the results show that there were not any significant differences between the two sets of scores at the level of .016 in the fluency measures of the number of words ( $t = -1.75$ ,  $df = 19$ ,  $sig > .016$ ), the number of T-units ( $t = -2.09$ ,  $sig > .016$ ), and the number of clauses ( $t = -2.60$ ,  $df = 19$ ,  $sig > .016$ ). Therefore, the results suggest that the learners in the peer-assisted group did not show any improvement in the fluency feature in the posttest.

Considering the complexity measures in the pretest and posttest, at the significance level of .025, there were not any significant differences between the two sets of scores in terms of clauses per T-unit ( $t = -.136$ ,  $df = 19$ ,  $sig > .025$ ) and dependent clauses percentage ( $t = -.250$ ,  $df = 19$ ,  $sig > .025$ ). Therefore, the learners in the peer-assisted group did not improve in terms of the complexity measures.

As for the peer-assisted group's performances in terms of the accuracy measures, there were significant differences between the two sets of scores in terms of error-free T-units percentage ( $t = -3.72$ ,  $df = 19$ ,  $sig < .025$ ) and error-free clauses percentage ( $t = -5.28$ ,  $df = 19$ ,  $sig < .025$ ). Therefore, the learners in the peer-assisted group did improve in terms of accuracy. Moreover, the eta squared statistic (.39 for error-free T-units percentage and .54 for error-free clauses percentage) indicates large effect sizes (Cohen, 1988).

**Table 3.** Paired-Samples T-Tests Comparing the Performances of the Peer-Assisted Group in the Pretest and the Posttest in Terms of the Subcategories of CAF Measures

	N	M	SD	T	df	Sig.
<b>Fluency measures</b>						
<i>Words</i>						
Pretest	20	158.16	36.45	-1.75	19	.096
Posttest	20	175.65	48.61			
<i>T-units</i>						
Pretest	20	10.95	2.72	-2.09	19	.050
Posttest	20	13.40	5.38			
<i>Clauses</i>						
Pretest	20	16.45	3.67	-2.60	19	.017
Posttest	20	19.65	6.18			
<b>Complexity measures</b>						
<i>Clauses per T-unit</i>						
Pretest	20	1.50	.24	-1.36	19	.893
Posttest	20	1.51	.28			
<i>Dependent clauses percentage</i>						
Pretest	20	33.57	10.71	-2.50	19	.806
Posttest	20	34.60	11.54			
<b>Accuracy measures</b>						
<i>Error-free T-units percentage</i>						
Pretest	20	60.57	18.39	-3.72	19	.001
Posttest	20	69.98	19.53			
<i>Error-free clauses percentage</i>						
Pretest	20	71.25	14.74	-5.28	19	.000
Posttest	20	79.82	14.69			

Table 4 represents the results of the t-tests which investigate the differences between the peer-assisted and AI-assisted groups' performances in the posttest in terms of CAF measures. Based on the Bonferroni adjustment, there were not any significant differences between the groups in terms of the fluency measures of the average words ( $t = .458$ ,  $df=38$ ,  $p > .016$ ), the number of T-units ( $t = 1.005$ ,  $df= 38$ ,  $p > .016$ ), and the number of clauses ( $t = 1.18$ ,  $df = 38$ ,  $p > .016$ ). Furthermore, with respect to the complexity measures, there were not any significant differences between the two groups in terms of clauses per T-units ( $t = -.440$ ,  $df=38$ ,  $p > .025$ ) and dependent clauses percentage ( $t=-.583$ ,  $df=38$ ,  $p > .025$ ). However, with regard to the accuracy, there were significant differences between the two groups in terms of error-free T-units percentage ( $t = -2.55$ ,  $df= 38$ ,  $p < .025$ ) and error-free-clauses percentage ( $t=-2.51$ ,  $df= 38$ ,  $p < .025$ ). Therefore, the learners in the peer-assisted and AI-assisted groups did not have similar performance in terms of the accuracy measures. As the learners' means in the accuracy measures show, the peer-assisted group outperformed the AI-assisted group in both of the accuracy measures. Moreover, the eta squared statistic (.34 for error-free T-units percentage and .36 for error-free clauses percentage) indicated large effect sizes (Cohen, 1988).

**Table 4.** Independent-Samples T-Tests Comparing the Peer-Assisted and AI-Assisted Groups in Terms of the Subcategories of CAF Measures in the Posttest

	Levene's test for equality of variances					t-test		
	N	M	SD	F.	Sig.	T	df	Sig.
<b>Fluency measures</b>								
<i>Words</i>								
Peer-assisted	20	182.45	45.29					
AI-assisted	20	175.65	48.61	.006	.937	.458	38	.650
<i>T-units</i>								
Peer-assisted	20	15.20	5.93					
AI-assisted	20	13.40	5.38	1.08	.305	1.005	38	.321
<i>Clauses</i>								
Peer-assisted	20	22.35	7.91					
AI-assisted	20	19.70	6.13	2.32	.136	1.18	38	.244
<b>Complexity measures</b>								
<i>Clauses per T-unit</i>								
Peer-assisted	20	1.49	.30					
AI-assisted	20	1.53	.27	.627	.433	-.440	38	.662
<i>Dependent clauses percentage</i>								
Peer-assisted	20	32.28	13.47					
AI-assisted	20	34.60	11.54	.633	.431	-.583	38	.563
<b>Accuracy measures</b>								
<i>Error-free T-units percentage</i>								
Peer-assisted	20	55.22	16.96					
AI-assisted	20	69.98	19.53	.841	.365	-2.55	38	.015
<i>Error-free clauses percentage</i>								
Peer-assisted	20	68.37	14.05					
AI-assisted	20	79.82	14.69	.244	.624	-2.51	38	.016

### Discussion of the Results of T-Tests Investigating Learners' Performances

The results showed that although the AI-assisted group improved in terms of fluency measures, the performance of this group in the pretest was not different from that in the posttest in terms of the complexity and accuracy features. In addition, while the peer-assisted group's performances did not differ in terms of complexity and fluency features between the pretest and posttest, they showed improvement in accuracy measures in their posttests.

Therefore, the learners in the peer-assisted group improved their performance just in terms of the accuracy measures, but not in the fluency and complexity measures. This finding is in line with the concept of limited attention resources advocated by Larsen-Freeman (2009) and Skehan (2009). According to these scholars, one cannot attend to all aspects of language, such as complexity, accuracy, and fluency, simultaneously. Therefore, a trade-off exists between these features. In other words, since the learners in the peer-assisted group focused on the accuracy of their productions, they neglected the complexity and fluency features.

In addition, the superior accuracy in the peer-assisted group compared to the AI-assisted group can be attributed to the unique advantages of human interaction in the collaborative writing process. First, compared to AI, peers may identify certain errors more successfully and provide higher-quality feedback in terms of clarity, accuracy, and critical aspects that may be

neglected by AI (Steiss et al., 2024). Second, peers will negotiate their thoughts, seek help, and assist each other due to the interactivity, which is absent in the AI-assisted group (Li, 2023). Third, when offering feedback, peers tend to focus on one aspect of performance instead of multiple aspects, which is visible in AI performance (Song & Song, 2023).

Importantly, while AI tools like ChatGPT are frequently seen as helpful for grammar, punctuation, and sentence corrections (surface-level errors), questions remain about the completeness, accuracy, and biases of the content-level suggestions made by AI tools. They are sometimes considered inaccurate or fake. Whereas, these types of errors can be detected by human more easily (Gayed et al., 2022; Wang, 2024). In addition, negotiation of meaning during peer interaction increases the learners' mutual understanding, decreases misinterpretations, and increases revisions in their texts (Miao et al., 2006).

Moreover, the more accurate performance of learners in the peer-assisted group can be due to the feedback they offered to each other and the feedback they received. This is in line with the results of the studies conducted on the effect of peer feedback on learners' performance. Peer comments are a valuable source of information. Even when learners are not proficient in second language writing, they can provide their peers with feedback on the drafts. Peers raise a sense of real audience in the learners; subsequently, they raise critical thinking and awareness of the strengths and weaknesses of their writings, and facilitate their collaborative endeavor (Xiaomeng & Ravindran, 2024; Zeng & Ravindran, 2025).

Furthermore, the results of this part of study are in line with those of McDonough (2004), who asserted that learners who engaged in negative feedback and modified output during collaborative tasks improved the accuracy of grammatical points they used. Similarly, the findings are consistent with those of Dobao's (2012) study, which revealed the positive impact of collaboration in pairs on the linguistic accuracy of the students' written products; grammatical and lexical accuracy of the texts written by the learners in pairs increased. In other words, the pair members collaboratively decided on the grammar and vocabulary of their production. In addition, Sang and Zou (2023), Chen (2019), and Storch (2005) found that the learners who worked collaboratively produced more accurate texts. Peer-assisted collaborative performance helped learners share their knowledge and give feedback to their partners.

However, the discrepancy between the results of the present study and those of Shehadeh's (2011) study concerning the accuracy of the learners' performance in the peer-assisted group may be due to the use of different measures of grammatical accuracy. In other words, in the current study, the proportion of error-free clauses and the proportion of error-free T-units were used as two measures of accuracy; however, in Shehadeh's (2011) study, global scales based on a rater's judgment of the text elements, such as content, organization, and grammar, were used as the indicator of accuracy.

In addition, the present study showed that collaborative performance did not affect the fluency of the learners' productions in the peer-assisted group. However, the fluency of the productions of learners in the AI-assisted group improved. The improved fluency of the AI-assisted group can be attributed to the AI tools' function to reduce the cognitive burden associated with tasks. By offloading the task, students are freed up and may have faster and more efficient writing process (Levine et al., 2024; Wang, 2024). Dhillon et al. (2024) have

also found that learners who are assisted with AI can lead to significant improvements in productivity (words/time) of texts.

Concerning the results about the learners' performance in terms of syntactic complexity, the lack of improvement in the AI-assisted group in terms of complexity can be attributed to the fact that AI may hinder students' creativity and critical thinking. While AI can reduce the cognitive load and pressure on tasks, it does not automatically result in complex texts. The use of extensive AI could reduce learners' efforts, creativity, and ability to produce complex texts (Nguyen et al., 2024; Song & Song, 2023; Wang, 2024).

### **Learners' Perceptions of their Writing Experiences**

When learners were asked about their views on the writing activity in this course, 75% of the learners in AI-assisted group had positive views of their writing experiences. Most of them were satisfied with the experience of learning academic writing using technology. For instance, Mehrdad stated that "I didn't know how to write a proper and standard text before participating in this writing class". However, 25% of the learners in AI-assisted group were not satisfied with their writing experiences. They claimed that lack of opportunity for real interaction made the experience boring. For instance, Maryam mentioned that "it was boring to write through asking the AI to answer my questions".

Concerning the learners in peer-assisted group, 90% of the learners had positive attitudes towards their writing experiences. They referred to the ease of generating ideas, the beneficial effect of receiving feedback during writing tasks, and the motivation gained by finding appropriate vocabulary and grammatical structures. For instance, according to Ali, "pair writing could help me brainstorm and find various ideas". However, 10% of learners did not have a positive attitude towards their writing experiences. According to Atefeh, "It was difficult to reach consensus on opinions and appropriate vocabulary".

When learners were asked about the most interesting part of the activity, most of the learners in the AI-assisted group (75%) believed that being corrected by the AI as soon as asking for feedback was interesting. For instance, Ali mentioned that "It was interesting for me to be corrected by ChatGPT".

Concerning the learners in peer-assisted group, most of them enjoyed the experience of collaborative brainstorming and offering and receiving feedback. For instance, Ehsan claimed that "finding ideas to write about is really hard for me but the collaborative tasks facilitated joint idea generation. I love this part the most".

When asked about the most difficult part of the activity, the learners in AI-assisted group referred to the poor internet quality, headache due to looking at the screen for a long time, being bored, and not being sure of the quality of information offered by the AI. According to Mahin, "sometimes, the net quality was poor which hindered using the AI appropriately".

However, the learners in peer-assisted group noted the inability to reach an agreement, dealing with an uncooperative partner, being forced to cooperate during the tasks, and experiencing a final test which was different from their class experience as the difficulties. For instance, Ali mentioned that "sometimes, my partner and I did not agree on the controlling ideas of the paragraphs".

When asked about their perceptions of writing before the experience, all learners in the AI-assisted group believed that their views on writing changed; they learned how to write academically using AI. For instance, Zahra mentioned that “at the beginning of the class, I didn’t know how to write in English; however, I am familiar with different methods of paragraph development now through writing using the AI”.

Among the learners whose views on writing changed in the peer-assisted group, five learners admitted that they used to think that there was just one mode of writing (i.e., individual writing). After the course, they were familiar with the learning potential of collaborative writing. For instance, Aziz stated that “I had thought that writing is limited to an individual endeavor. However, the presence of a partner to assist in writing was a great idea”. Moreover, four learners mentioned that they realized that writing was easier than what they had thought. For instance, Leili stated that “I previously thought that I could never write a standard text in English. But my partner motivated me and provided feedback which made the attempt easier”. However, five learners did not change their minds and regarded collaborative writing an interesting experience.

When asked about the possible effects of the activities on their confidence in their writing ability, 60% of the learners in AI-assisted group believed that their writing experience had beneficial effects on improving their confidence in writing. For instance, Ali mentioned that “I am self-confident since I know how to write a good paragraph using AI”. 50% of the learners referred to their satisfaction with detecting their problematic areas to overcome them. Nikan admitted that “when the AI helped me to detect my problems in writing, I did my best to correct the errors in my writings”. However, 40% of learners admitted that they completely lost their self-confidence during writing tasks. Mohsen mentioned that “I lost my self-confidence because I realized that there was an error in every sentence which I wrote”.

In addition, among the learners in peer-assisted group, 85% of the students expressed the beneficial effect of collaborative writing. More specifically, they referred to the benefits of collaborative writing on improving the quality of compositions, gaining confidence in writing, and receiving feedback and affective support from their partners during collaborative writing. According to Mina, “I have a problem writing alone. I really need someone to help me and provide me with the feedback”. Conversely, 15% of the learners believed that collaborative writing had negative effects on them. Sina stated that “I feel the pressure of the must to cooperate in the class and be active. This hinders me from writing well”.

When asked if the activity had any effect on their other language skills, 60% of the learners in AI-assisted group admitted the beneficial effects of the activities on improving reading, vocabulary, and grammar knowledge. Manije admitted that “I focused on the methods of paragraph development in the paragraphs offered by the AI”. However, 40% of the learners considered the activities ineffective in improving other language skills. For instance, Soheil believed that “I couldn’t grasp the relationship of writing skill using AI with other skills”.

In addition, among the learners in peer-assisted group, although 10% of the learners could not find a relationship between writing and other skills, 90% of the learners admitted to the beneficial effects of the activities on improving speaking, listening, reading, and vocabulary areas. For instance, Ahmad noted that “I think collaborative writing helped me speak more

fluently”.

Finally, when asked if they want to continue participating in the writing activities they did in the class, 65% of learners in AI-assisted group and 80% of learners in peer-assisted group preferred to have similar activities in the future.

### **Discussion on Learners' Perceptions of their Writing Experiences**

The study revealed that while learners expressed satisfaction with both AI-assisted and peer-assisted writing experiences, the peer-assisted group reported a higher level of satisfaction. Actually, the learners who wrote in the AI-assisted group referred to some problems that were solved for the peer-assisted writing group. For instance, the learners in the AI-assisted group mentioned the lack of real interaction, interruptions caused by poor internet quality, and physical problems (headache) caused by staring at the screen.

Conversely, 80% of learners in the peer-assisted group had overall positive attitudes toward pair work. The satisfaction expressed by the majority of the learners with their peers' assistance and collaboration is in line with the findings of many studies (e.g., Ajmi & Ali, 2014; Alqasham, 2022; Dobao & Blum, 2014; Le et al., 2018; McKay & Sridharan, 2024; Shehadeh, 2011).

Furthermore, most of the learners in the peer-assisted group admitted that their collaborative writing experience provided them with the opportunity to cooperate, share their ideas, and find appropriate words and structures. This perception reflects the findings of previous research focusing on the merits of peer-assisted collaborative tasks (e.g., Bueno-Alastuey et al., 2022; Dobao, 2012; Pham, 2021; Sang & Zou, 2023). The beneficial effect of peer-assisted writing is due to the interaction of learners and contextual understanding (Escalante et al., 2023) which facilitates the possibility of asking follow-up questions (Li, 2023), providing affective feedback (Escalante et al., 2023), and retaining their personal voice (Song & Song, 2023).

Moreover, contrary to the learners in the AI-assisted group, most of the learners in the peer-assisted group believed that the collaborative writing tasks positively affected their speaking and listening skills. This perception is in line with the principles of the sociocultural theory which suggests that “learning and development is the socio-genesis product of meaningful social interactions among the community members in the respective learning context” (Ahangari et al., 2014, 84). Therefore, instead of believing in the universal features of cognition, the sociocultural theory proposes that learning emerges as a result of the learners' interaction in social and cultural contexts (Johnson, 2009).

Contrary to the AI-assisted group, most of the learners in the peer-assisted group referred to the difficulty in agreeing with their partners. Different factors might have affected this dissatisfaction, including the learners' patterns of dyadic interaction (Storch & Aldosari, 2012; Watanabe, 2008), learners' unequal responsibility in group activities (Le et al., 2018), and differences in the gender of the learners in a pair (Bacon, 2005). However, research findings show that learners will gradually learn how to collaborate with their partners and how to deal with disagreements during collaborative activities (Sridharan & Boud, 2019).

On the other hand, AI-assisted group's lower satisfaction with their experience can be attributed to the decrease in the students' perceived ownership of the text (Dhillon et al., 2024),

limitation in students' creativity and critical thinking (Escalante, et al., 2023; Malik et al., 2023), mistrust to the accuracy of information provided by the AI (Levine et al., 2024), and the lack of a synchronous communication (Li, 2023).

## Conclusion

This mixed-method study deeply examined students' writing endeavors in peer-assisted and AI-assisted writing groups, taking into account their performances and perceptions. The study showed that although the learners in both groups had similar performances in terms of fluency and complexity measures, the learners in the peer-assisted group outperformed in terms of accuracy measures. In addition, the coding of the learners' views showed that the learners in the peer-assisted group were generally more satisfied with their writing experiences. Moreover, the learners writing in the AI-assisted group reported some difficulties and problems, which were not reported by the learners in the AI-assisted group.

While this study compared peer-assisted and AI-assisted collaborative writing, it is important to acknowledge that the two conditions involved different forms of interaction. Peer-assisted writing reflected a socio-collaborative process, characterized by mutual negotiation, shared metacognition, and co-regulation between human partners. In contrast, the AI-assisted condition represented a human–tool interaction, where learners engaged in guided prompting, reflection, and selective adoption of AI-generated input. Therefore, the study compared not merely the *source* of assistance (peer vs. AI), but also the nature of collaboration (human-to-human versus human–AI co-construction). Consequently, observed differences in writing outcomes and perceptions may reflect both the type of assistance and the interactional dynamics embedded in each condition. We thus interpret our findings cautiously, positioning AI as a cognitive and linguistic scaffold that complements the socio-collaborative benefits of peer interaction.

These findings have significant implications for curriculum designers and teachers. Due to the beneficial effects of collaborative tasks, collaborative activities (specifically peer-assisted ones) should be designed in the curriculum. In addition, being aware of the usefulness of collaborative endeavors and students' positive perceptions of collaboration in both groups, teachers can assign students to groups and conduct collaborative writing tasks to motivate them to learn in a friendly context, be interactive, and have a voice to feel less bored in the class. They can also introduce AI to the students to benefit from its capabilities.

This study includes some limitations. First, the sample size was limited; further studies can replicate this investigation and examine a larger sample size. Second, the students' performances were examined cross-sectionally. Thus, future research can examine students' performances at different data points during the semester. Third, this study was conducted using one specific type of AI (ChatGPT). Thus, further studies can use other AI to examine its possible effects to improve the learners' productions. Forth, a delayed post-test to assess the retention of writing gains was absent. Future studies should include delayed assessments to determine whether the observed gains are sustained over time.

## References

Ahangari, S., Hejazi, M. & Razmjou, L. (2014). The impact of scaffolding on content retention of Iranian post-elementary EFL learners' summary writing. *Procedia—Social and Behavioral Sciences*, 98, 83-89. <https://doi.org/10.1016/j.sbspro.2014.03.392>

Ajmi, A. A. S. & Ali, H. E. (2014). Collaborative writing in group assignments in an EFL/ESL classroom. *English Linguistics Research*, 3(2), 3-17. <https://doi.org/10.5430/elr.v3n2p1>

Alqasham, F. (2022). Investigating English as a foreign language learners' perceptions, emotions, and performance during online collaborative writing. *Frontiers in Psychology*, 13, 954011. <https://doi.org/10.3389/fpsyg.2022.954011>

Anshu, A., H. & Yesuf, M., Y. (2022). Effects of Collaborative Writing on EFL Students' Paragraph Writing Performance: Focus on Content and Coherence. *International Journal of Education and Literacy Studies*, 10(1), 36-47. <https://doi.org/10.7575/aiac.ijels.v.10n.1p.36>

Ary, D., Jacobs, L. C. & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). Thomson.

Bacon, D. R. (2005). The effect of group projects on content-related learning. *Journal of Management Education*, 29(2), 248-267.

Bueno-Alastuey, C., Vasseur, R. & Elola, I. (2022). Effects of collaborative writing and peer feedback on Spanish as a foreign language writing performance. *Foreign Language Annals*, 1-23. <https://doi.org/10.1111/flan.12611>

Chen, W. (2019). An exploratory study on the role of L2 collaborative writing on learners' subsequent individually composed texts. *Asia-Pacific Educational Research*, 1-11. <https://doi.org/10.1007/s40299-019-00455-3>

Cohen, J.W. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> edn). Lawrence Erlbaum Associates.

Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping human-AI collaboration: Varied scaffolding levels in co-writing with language models. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1-18. <https://doi.org/10.1145/3613904.3642134>

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers and Education*, 210, 104967.

Dobao, A., F. & Blum, A. (2014). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, 41, 365-378. <https://doi.org/10.1016/j.system.2013.02.002>

Escalante, J. & Pack, A. & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and student preference. *International Journal of Educational Technology in Higher Education*, 20(57), 1-20. <https://doi.org/10.1186/s41239-023-00425-2>

Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299-323. <https://doi.org/10.1017/S0272263100015047>

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055.

Hiromori, T. (2021). Are two heads better than one? Comparing engagement between pairs and

individuals in an L2 writing task. *Journal Language Teaching Research Quarterly*, 21, 66-83. <https://doi.org/10.32038/ltrq.2021.21.05>

Hsu, H. (2025). The effect of collaborative prewriting on L2 collaborative writing production and individual L2 writing development. *International Review of Applied Linguistics in Language Teaching*, 63(1), 499-533. <https://doi.org/10.1515/iral20230043>

Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory into Practice*, 38, 67-73. <https://doi.org/10.1080/00405849909543834>

Johnson, K., E. (2009). *Second language teacher education: A sociocultural perspective*. Routledge.

Kim, M., Kim, N. & Heidari, A. (2022). Learner experience in artificial intelligence-scaffolded argumentation. *Assessment and Evaluation in Higher Education*, 8(47), 1301-1316. <https://doi.org/10.1080/02602938.2022.2042792>

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579-589. <https://doi.org/10.1093/applin/amp043>

Le, H., Janssen, J. & Wubbels, T. (2018). Collaborative learning practices: Teacher and student perceived obstacles to effective student collaboration. *Cambridge Journal of Education*, 48(1), 103-122. <https://doi.org/10.1080/0305764X.2016.1259389>

Levine, S., Beck, S. W., Mah, C., Phalen, L., & PIltman, J. (2024). How do students use ChatGPT as a writing support? *Journal of Adolescent and Adult Literacy*, 68(5), 445-457. <https://doi.org/10.1002/jaal.1373>

Li, A. W. (2023). Using Peerceptiv to support AI-based online writing assessment across the disciplines. *Assessing Writing*, 57, 100746.

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, 5, 100296.

McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32, 207-224. <https://doi.org/10.1016/j.system.2004.01.003>

McKay, J. & Sridharan, B. (2024). Student perceptions of collaborative group work in higher education. *Studies in Higher Education*, 49(2), 221-234. <https://doi.org/10.1080/03075079.2023.2227677>

Mei, F., Ma, Q., Tang, J., & Zou, B. (2024). Exploring collaborative writing among large groups in online distance learning through a public forum and private chat tool. *Journal ReCALL*, 37(1), 1-19. <https://doi.org/10.1017/S0958344024000211>

Miaoa, Y., Badger, R. & Zhen, Y. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179-200.

Moxie, J., Dahl, A. A., Fernandez-Borunda, A., & Partridge, H. (2025). Teamwork makes the dream work: Group effectiveness in a paper chase collaborative writing exercise for higher education. *Frontiers in Education*, 10, 1-9.

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847-864. <https://doi.org/10.1080/03075079.2024.2323593>

Norris, J. M. & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed

SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578. <https://doi.org/10.1093/applin/amp044>

Özçelik, N. & Yangın Ekşi, G. (2024). Cultivating writing skills: The role of ChatGPT as a learning assistant: A case study. *Smart Learning Environments*, 11, 1-18. <https://doi.org/10.1186/s4056102400296-8>

Pandemic. (n.d.) *Foreign Language Annals*, 54(3), 714-739.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315-341. <https://doi.org/10.1007/s10648-006-9029-9>

Pekrun, R. (2018). Control-value theory: A social-cognitive approach to achievement emotions. In G. A. D. Liem, & D. M. McInerney (Eds.), *Big theories revisited 2: A volume of research on sociocultural influences on motivation and learning* (pp. 162-190). Information Age Publishing.

Peng, Y., Li, Y., Sy, Y., Chen, K. & Jiang, S. (2022). Effects of group awareness tools on students' engagement, performance, and perceptions in online collaborative writing: Intergroup information matters. *The Internet and Higher Education*, 53, 100845. <https://doi.org/10.1016/j.iheduc.2022.100845>

Pham, V., P. (2021). The effects of collaborative writing on students' writing fluency: An efficient framework for collaborative writing. *Sage Open*, 1-11.

Qiu, X., & Lee, M. K. (2020). Regulated learning and self-efficacy beliefs in peer collaborative writing: An exploratory study of L2 learners' written products, task discussions, and self-reports. *System*, 93, 102312. <https://doi.org/10.1016/j.system.2020.102312>

Sang, Z. & Zou, W. (2023). The effect of joint production on the accuracy and complexity of second language writing. *Journal of Psycholinguistic Research*, 52, 425-443. <https://doi.org/10.1007/s10936-022-09882-8>

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20, 286-305. <https://doi.org/10.1016/j.jslw.2011.05.010>

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532. <https://doi.org/10.1093/applin/amp047>

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, 1-14.

Sridharan, B., Tai, J. & Boud, D. (2019). Does the use of summative peer assessment in collaborative group work inhibit good judgments? *Higher Education* 77(5), 853-870. <https://doi.org/10.1007/s10734-018-0305-7>

Sridharan, S., Muttakin, M., B. & Mihret, D.G. (2018). Students' perceptions of peer assessment effectiveness: An explorative study. *Accounting Education* 27(3), 259-85. <https://doi.org/10.1080/09639284.2018.1476894>

Steiss, J., Tate, T. P., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. (2024). Comparing the quality of human and ChatGPT feedback on students' writing. *Learning and Instruction*, 91, 101894.

Storch, N. & Aldosari, A. (2012). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31-48. <https://doi.org/10.1177/1362168812457530>

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

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.

Styati, E. W. (2018). Investigating dominant and passive students on pair work towards the students' writing performance, *3L: Language, Linguistics, Literature*. 24(3), 142-154. <https://doi.org/10.17576/3L-2018-2403-11>

Su, F., Zou, D. & Xie, H. (2024). Integrating different group patterns into collaborative argumentative writing in the Shimo platform. *Journal Technology, Knowledge and Learning*, 29(1), 309-330. <https://doi.org/10.1007/s10758-022-09632-8>

Svalberg, A. M.-L. (2009). Engagement with language: interrogating a construct. *Language Awareness*, 18, 242–258. <https://doi.org/10.1080/09658410903197264>

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning*, 30, 1-22. <https://doi.org/10.1007/s10758-024-09744-3>

Watanabe, Y. (2008). Peer-peer interaction between L2 learners of different proficiency levels: Their interactions and reflections. *The Canadian Modern Language Review*, 64(4), 605–635. <https://doi.org/10.3138/cmlr.64.4.605>

Wigglesworth, G. & Storch, N. (2012). What role for collaboration in writing and writing feedback? *Journal of Second Language Writing*, 21, 364–374.

Wolfe-Quintero, K., Inagaki, S. & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawaii Press.

Xiaomeng, Z. & Ravindran, L. (2024). The effects of peer feedback provision and reception on lower-proficiency EFL learners' writing development. *Studies in Educational Evaluation*, 83, 101422. <https://doi.org/10.1016/j.stueduc.2024.101422>

Zeng, X. & Ravindran, L. (2025). Design, implementation, and evaluation of peer feedback to develop students' critical thinking: A systematic review from 2010 to 2023. *Thinking Skills and Creativity*, 55, 101691. <https://doi.org/10.1016/j.tsc.2024.101691>

Zhang, M. (2019). A re-examination of pair dynamics and L2 learning opportunities in collaborative writing. *Language Teaching Research*, 23(6), 683-804.

Zhang, M. (2022). A re-examination of pair dynamics and L2 learning opportunities in collaborative writing. *Journal Language Teaching Research*, 26(1), 10-33. <https://doi.org/10.1177/1362168819890949>

## Appendix A

The prompt that learners were provided with to give to the AI for collaborative writing was as follows:

*You are my collaborative writing partner. We are two EFL students working together to write a paragraph on “[insert topic].” Please help me through all stages of the writing process (brainstorming, drafting, and revising/editing) as if you were an equal team member, not a teacher or editor.*

### Here's what we would like you to do:

1. **Brainstorming:** Ask me questions to help us clarify my position and generate main ideas, supporting details, and counterarguments. Don't give full answers immediately; help me think by prompting and discussing ideas together.
2. **Drafting:** Work with me to organize the essay and develop each paragraph. When suggesting sentences or phrases, explain *why* you think they fit. Let's negotiate wording and structure collaboratively.
3. **Revising and editing:** Read my draft carefully and provide feedback on content, coherence, vocabulary, and grammar. Suggest improvements, but ask for my opinions before finalizing changes.

**Output format:** Please divide your responses into the three stages (Brainstorming / Drafting / Revising) and always invite my input before moving forward. Let's build the paragraph together step by step.

## Appendix B

### An example of student–AI interactions

**Student A:** I want to write a paragraph about the impact of social media on student learning. Can you help me get started?

**AI:** Of course! To begin, do you want to focus on the positive effects, the negative effects, or both?

**Student A:** Maybe both of them, but focus on the positive effects first.

**AI:** Great. Let's generate some ideas for benefits. How about thinking of ways social media can help with collaboration, motivation, or access to information? Which of these seems most relevant to your experience?

**Student A:** I think collaboration fits best because I often use online groups to study together.

**AI:** Excellent. Now, what evidence or examples could you include to show how collaboration improves learning?

### An example of peer–peer interactions

**Student B:** We have to write a paragraph about social media and student learning. Is it good to write an outline for this topic?

**Student C:** Yes, think about the major supports first.

**Student B:** What are the different effects of social media on learning?

**Student C:** It makes the process easier. Let's write... social media make learning easier.

**Student A:** No, makes.

## Appendix C

### **Shehadeh's (2011) Survey of Students' Views and Perceptions of Writing Experience**

Dear Students,

Thank you for participating in this study.

In order to better understand the usefulness of your writing experience, we would like you to answer the following questions based on your views, perceptions, and experience of the writing activity used in this course. You are strongly encouraged to provide your opinions frankly. Your answers will remain anonymous to everyone.

1. What's your view on the writing activity in this course? How did you find the experience?
2. What was the most interesting part of the activity, or the part you liked most? Why?
3. What was the most difficult part of the activity? Why?
4. What was your view and perception of writing before the experience? Did these change after the experience?
5. Did the activity have any effect (positive or negative) on your confidence in your writing ability? Please specify.
6. Did the activity have any effect (positive or negative) on your other language skills (e.g., speaking, reading, and listening)? Please specify.
7. Would you like to do similar writing activities/tasks in the future?