

AUTOMATED WRITING EVALUATION (AWE) IN VIETNAMESE EFL CLASSROOMS:

ENHANCING SKILLS AND UNDERSTANDING STUDENT PERCEPTION

LE THI THIEN PHUOC¹, NGUYEN LAM ANH DUONG¹, PHAM THI THU¹, DANG HOANG MAI¹,
NGUYEN XUAN HONG*²

¹ Faculty of Foreign Languages – Industrial University of Ho Chi Minh City

² Industrial University of Ho Chi Minh City

*Corresponding author: nguyensexuanhong@iuh.edu.vn

DOIs: <https://www.doi.org/10.46242/jstiuh.v73i1.5095>

Abstract. This study aims to investigate the effectiveness of an Automated Writing Evaluation (AWE) tool, - *Write & Improve* - in enhancing writing skills among a group of students who are studying English as a foreign language at a public university in Vietnam and examine their perceptions of the advantages and disadvantages of this tool. The study participants included thirty-eight students taking a General English 1 course. A mixed methods design was employed for the study, which involved data collection through pre- and post-tests and a questionnaire on students' perception of the tool after one semester of use. This study's findings proved this tool's effectiveness in improving students' writing skills. They highlighted students' positive perceptions towards implementing the AWE tool in English as a Foreign Language (EFL) writing classrooms. However, the data also indicated that the students encountered several challenges. Teachers must consider these challenges when incorporating this tool into EFL writing instruction. Recommendations and practical implications for language teaching, learning, and assessment are discussed.

Keywords. AI, Artificial Intelligence, automated writing evaluation, AWE, EFL writing instruction, Write & Improve

1 INTRODUCTION

Artificial intelligence (AI) tools' increasing importance and popularity require teachers and students to make essential changes from the conventional teaching and learning style to a more student-centred and flexible style by integrating AI tools.

Among basic communication skills, English writing skills have become an urgent requirement for students while studying and become even more necessary after graduation. Students need this skill for basic tasks such as writing emails or CVs. However, writing is still challenging for most English-as-a-foreign/second language (EFL/L2) students. As Campbell (2019) stated, writing in English is a complicated, necessary, and integrative endeavour that both native and international students struggle with. In the context of public universities in Vietnam, EFL students usually have little time to practice English writing in class and typically do not receive frequent feedback on their English writing from the lecturer due to large classes. However, giving detailed and timely feedback to many students is challenging for the teacher in charge, and delays and disruptions could arise (Boud & Molloy, 2013). Besides, it takes teachers too much time and effort and is subjective to monitor the writing process and provide relevant, helpful feedback to students (Lim and Phua, 2019). Therefore, AI-based AWE tools can be a potential solution for students and teachers since they provide extra writing practice with various topics and give corrective feedback and an overall student score.

Several studies have proven the positive relationship between the feedback provided by AWE tools and the improvement of students' academic performance in writing skills and their general foreign language skills (Elliot & Mikulua, 2004; Shermis et al., 2008). Lim and Phua (2019) also highlighted the advantages of incorporating technology in classrooms, particularly in offering instant feedback and generating ideas during the drafting and revision phases of process writing, which might enable immediate adjustments to language usage. As a result, this would foster self-editing and improvements among learners and reduce the waiting time for instructors' feedback and corrections.

Considering these factors, this study focuses on exploring the efficacy and implementation of integrating AWE tools into the teaching and learning of writing skills for non-English-majored students at the university level. Additionally, a survey was conducted on the students' perceptions of the tool's advantages and disadvantages to gain valuable insights. By delving into theoretical frameworks and practical implementation strategies, this research aims to provide insights into the possibilities of incorporating *Write & Improve* - an AWE tool- to enhance non-English majors' writing skills and reduce teachers' workload in large classes. Through meticulous examination and critical analysis, this research will pave the way for more effective technology integration in language education.

2 LITERATURE REVIEW

2.1 Automated Writing Evaluation Tools

Many technology companies or institutions have developed AWE programs with various names, purposes, and features, such as *Grammarly*, *Jukuu*, *My Access*, *Criterion*, or *Write and Improve*. Generally, AWE is often called an automated writing feedback program that delivers suggestive and corrective feedback to a text (Ariyanto et al., 2021).

Automated writing evaluation tools can evaluate writing by comparing learners' writing with an extensive database of samples of the same genre. From there, they can score student's writing or arrange their writing into a certain level according to the Common European Framework of Reference (CEFR) competency framework. AI-based AWE tools can provide real-time feedback that adjusts student progress (Zhu et al., 2020; Ranalli et al. (2017).

2.2 Write & Improve - an AWE tool

Write & Improve is a free-of-charge AWE tool that uses new technology to develop for learners of English that marks writing in seconds. The tool was developed at the University of Cambridge to assess learners' English writing proficiency. When students submit their work, *Write & Improve* will score it on the CEFR scale, giving it a level from A1 (lowest) to C2 (highest). This tool also shows the writer the parts of the text that may need improvement. By doing so, students can work more on these areas and keep improving.

2.3 The integration of AWE tools into the teaching, learning and assessment of EFL writing skill classrooms

Existent literature points out that different applications of AWE have been used in teaching and learning writing in recent years. The studies indicated the benefits brought back by these tools, for instance, in a study with a combination of interviews and surveys conducted by Li, Link, and Hegelheimer (2015) to determine the impact of *Criterion* on students' writing ability and the teaching of writing skills. The results showed that the tool helped students make more edits, and the accuracy from the draft to the final version improved due to the comments given by the tool. Similarly, Marzuki et al. (2023) also surveyed teachers' views on the impact of AI tools on the content and layout of students' writing with a qualitative method through interviews with 4 English teachers at three different universities in Indonesia. The results also show that teachers use *Quillbot*, *Wordtune*, *Jenni* or *GPT Chat tools*. The participating teachers agreed that these AI tools help increase the quality of writing, especially the quality of content and structure. In the same year, Alharbi (2023) conducted a study about how AI tools support learners' writing skills. The study indicated that AI writing support technologies, which produce writing very similar to human writers, will be a potential solution in supporting learners' writing skills, primarily academic writing. The studies prove that AI tools are effective in improving writing skills, helping to make learning to write more exciting and assisting students to realize the importance of learning (Abalkader, 2022).

In terms of assessment, *Write and Improve*, as well as other AWE and scoring tools such as *Criterion*, are mentioned in several articles as tools with diverse applications that can be used in teaching and learning and, testing and assessing writing skills. For example, Hockly (2018) mentions that these tools can be used for placement tests, formative assessments, and end-of-course assessments. These tools provide quick feedback and have an interactive interface, making them convenient for evaluating writing for many students. However, research indicates inconsistency between scores given by teachers and these tools (Elliot et al.2013). In 2020, Karpova conducted a study on the application of *Write & Improve* in the context of university-level English teaching in Ukraine. The study participants were a small group of 12 English

majors with proficiency levels ranging from B1 to B2. The results showed that this tool is highly effective and valuable for teaching and assessment.

Providing timely feedback helps students take charge of their learning and aligns with the principles of assessment for learning (Black & Wiliam, 1998). Incorporating technology in classrooms, particularly in terms of offering instant feedback and generating ideas during the drafting and revision phases of process writing, enables learners to make immediate adjustments while learning the language (Lim & Phua, 2019) because this would foster self-editing and improvements among learners, as well as reducing the waiting time for instructors' feedback and corrections.

2.4 Learners' and teachers' perception of AWE tools

Regarding the learner's perspective, in 2019, O'Neill and Russell conducted a study on *Grammarly*, examining students' perceptions of using the application and teacher guidance. Their results showed that students who received feedback from both *Grammarly* and the instructors expressed a higher satisfaction level with the grammar instructions delivered, in contrast to those who obtained sole input from the instructors. The researchers proposed that such joined endeavours could supplement teachers' support to students and help alleviate issues related to inadequate time to address grammatical errors in students' essays. Miranty et al. (2021) conducted a study on the perceptions of university students in Indonesia on the advantages and disadvantages of the *Grammarly* tool in learning writing skills. Through a survey of 100 students at three levels: first year, second year, and third year, the results show that most students think that using *Grammarly* is very necessary to correct their writing since they get immediate comments, feedback and suggestions on improvement. Besides, the tool is easy to use and valuable. However, students at different levels commented differently on the usefulness of the comments provided by *Grammarly*. In 2022, Burkhard conducted a quantitative study exploring students' attitudes and perceptions about AI tools for writing skills at a university in New Zealand with 365 first-year students. The results show that students have different attitudes about AI-powered tools. Many students use these tools without critical thinking, just as a quick answer tool, which can lead to plagiarism. There are also many students complaining about unfairness when there are students who can use this tool to get good scores while other students cannot afford to use paid tools. Some students know how to use information selectively, but these students do not have an effective learning method. Therefore, the study suggests that teachers need to have more personalized teaching strategies for different groups of students so that students can see the effectiveness of these tools and be more careful when using them.

Putra, M. (2023) recently conducted another quantitative study on teachers' and students' perceptions of AI tools used in writing skills in teaching and learning. The study involved the participation of 18 lecturers and 39 graduate students in a TESOL course in the UK. The results are positive, as most teachers and students think using AI tools to correct writing errors is beneficial and does not violate academic integrity. Most believe that AI tools should be used to teach writing skills.

The above studies have shown the diverse applications of AWE tools in teaching writing skills and the attitudes and opinions of teachers and students about applying these technologies. Most of these studies survey the attitudes and perceptions of teachers and learners, and the results show that teachers and students have quite positive attitudes about applying AWE in learning and teaching writing. They also believe these tools can improve the quality of students' writing. However, little experimental research has shown whether AWE improves learners' writing skills after a course. A few experimental studies explored the effectiveness of AWE tools. Still, the number of participants is small, like in Karpova's (2020) study, which had only 33 participants or only four teachers, as in the study by Marzuki et al. (2023). Besides, no studies have been recorded on applying the Write and Improve tool to improve students' writing skills in Vietnam. Therefore, a study exploring the impact of AWE on students' writing performance and their perception towards using this tool needs to be implemented to figure out the effectiveness of this tool for large classes in the context of public universities in Vietnam.

The research is aimed to answer the following research questions (RQ):

*RQ1: How does **Write & Improve** enhance EFL students' writing skills?*

*RQ2: What are the advantages and challenges of utilizing **Write & Improve** in learning writing from the EFL students' perspectives?*

3 RESEARCH DESIGN

3.1 Research design

This research aimed to employ a mixed methods approach to recognize the effect of using Write and Improve in improving writing skills and perceived benefits and concerns towards this implication. Students' perceptions could be collected based on their responses on questionnaires, including closed and open-ended questions. Data obtained from pre and post-writing tests using Write and Improve were analyzed in an appropriate IBM statistical package and service solution (SPSS 22) test to test how significant the application is as treatment in this research could noticeably enhance students' writing skills. By combining both quantitative and qualitative methods, this mixed-method approach provides a comprehensive understanding of how "Write and Improve" influences students' writing skills from a statistical and experiential perspective.

3.2 Demographic Data of the Participants

The study participants included 38 students enrolled in a "General English 1" course in three months. Their English proficiency was at the A2 level (CEFR). The participants were assigned tasks as homework to use the *Write & Improve* tool. Regarding gender identity, there were 26 female and 12 male students, making up 68% and 32% among the overall sample, respectively.

3.3 Research instrument

Quantitative data

The quantitative aspect of the research focuses on measuring the improvement in students' writing skills before and after using "Write and Improve." A pre-test and post-test design is employed, where students take a writing test before using the tool (pre-test) and another after a set period (post-test). The hypothesis testing is listed out as follows:

One/ Null Hypothesis (H₀): There is no significant improvement in students' writing skills after using "Write and Improve."

Two/ Alternative Hypothesis (H₁): There is a significant improvement in students' writing skills after using "Write and Improve."

If the p-value of the t-test is less than 0.05, the null hypothesis is rejected, indicating a statistically significant improvement in writing skills. To assess this value, a scale shall be needed, whereby a rubric was used to measure various aspects of writing such as grammar, vocabulary, coherence, and overall structure, with consistent scoring criteria applied across both the pre-test and post-test.

Assessment Scales used for the assessment of the Pre-test and Post-test

Source: Cambridge English A2 Key 2020, Sample Tests Reading and Writing

Assessment Scales

Answers for A2 Writing tasks are assessed using the assessment scales, which consist of three subscales. Content, Language and Organization. A candidate can get up to 5 points for each criterion, which gives a total of 15 per task, in total, there are 30 possible marks in the Writing paper. Here is what the assessment criteria look like:

Table 1: The assessment criteria

Band	Content	Organization	Language
5	All content is relevant to the task Target reader is fully informed	Text is connected and coherent, using basic linking words and a limited number of cohesive devices	Uses everyday vocabulary generally appropriately, while occasionally overusing certain lexis Uses simple grammatical forms with a good degree of control While errors are noticeable, meaning can still be determined
4	Performance shares features of Bands 3 and 5		

3	Minor irrelevances and Omissions may be present The target reader is overall informed	Text is connected using basic, High frequency linking words	Uses basic vocabulary reasonably Appropriately Uses simple grammatical forms with some degree of control Errors may impede meaning at times
2	Performance shares features of Bands 1 and 3		
1	Irrelevances and misinterpretation of tasks may be present The target reader is minimally Informed	Production unlikely to be connected, though punctuation and simple connectors (i.e. "and") may, on occasion be used	Produces basic vocabulary of isolated words and phrases Produces few simple grammatical forms with only limited control
0	Content is irrelevant. The target reader is not informed	Performance below Band 1	

3.4 The experimental process & data collection

The experimental process spanned 13 weeks, starting with a paper-based pre-test in week 1, where the participants were asked to write an event or a story. Two independent examiners marked the pre-test papers. The pre-test scores were then converted to the A1-C1 scale. From week 2 to week 12, the students used the *Write & Improve* tool to practice their writing skills. The participants were informed why and how to use the tool effectively. They were asked to capture evidence of their progress and submit the screenshots to their portfolios. The weekly writing tasks were writing short stories about a past event, a camping trip, a meeting with someone they admire or idolize, a risk they have experienced, etc. All tasks were aligned with the lessons learnt in class during the course. In the Portfolio, students needed to submit:

- (1) their original writings
- (2) the improved version suggested by the tool, including the assessment of the relevance to the topic assigned and the level acquired
- (3) students' reflection, or a summary, on what they have learnt from the suggested version
- (4) screenshots of the relevance level to the topic (on a scale of 1 to 5), &
- (5) the level that the tool assessed.

During those ten weeks of working on this AWE tool, the teacher researchers frequently provided feedback and comments on the students' works. In Week 13, the students completed their final writing assignment as the post-test of the study. The participants were required to write a story about an event, with the topic being the same as the pre-test assignment. The post-test was conducted using a paper-based format. Two independent teachers who assessed the pre-test papers accessed all the post-test papers. The post-test scores were then converted to the A1- C1 scale. After completing the posttest, a questionnaire was administered to gather students' perceptions of the advantages and disadvantages of the tool, probing their views on its benefits and challenges.

3.5 Data analysis

With the pre-test and post-test scores, a paired-sample t-test was employed to compare the students' writing proficiency before and after the intervention. The paired-sample t-test was used in this study to confirm the alternative hypothesis that students exhibited improved performance after the ten weeks of practice writing with the AWE tool, with a statistically significant difference observed. Visual distributions of the scores at two different times were presented.

Additionally, numeric and textual data obtained from the questionnaire administered over the thirteen-week period were converted into percentages and mean scores to analyse students' perception.

4 FINDINGS & DISCUSSION

4.1 The enhancement of writing skills using *Write & Improve* on students' test scores

The rubric assessing paper-based students' writing on pre-tests and post-tests on *Write & Improve* evaluates three critical criteria: content, organization, and language. Each criterion is rated on a scale from 0 to 5, with a cumulative maximum of 15 points. To be more specific, “content” measures relevance, clarity, and task completion, while “organization” tests ideas' logical flow and coherence. Finally, “language” examines vocabulary usage, grammatical accuracy, and stylistic appropriateness.

For data analysis, the pre- and post-test scores were initially imported to conduct a normality test using the Shapiro-Wilk formula within the SPSS 22 program to measure their normality distribution. The obtained p-value indicates that the data were normally distributed, as presented in Table 2 below:

Table 2: Normality Test Results

Data	Shapiro-Wilk		
	Statistic	Df	Sig.
Pre-test	.947	38	.072
Post-test	.963	38	.238

The descriptive results of participants' performance on the pretest and posttest are shown in Table 3, in which the scores range from 0 to 12 with a mean of 4.842 in the pretest, while the mean was significantly higher for the posttest ($M = 9.763$), with the scores ranging from 3 to 15. Skewness values were positive with pre-test and post-test, showing 0.189 and 0.082, respectively.

Table 3: Descriptive statistics for pre and post-tests

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness
Pre-test	38	1.0	9.0	4.842	1.9105	.189
Post-test	38	3.0	15.0	9.763	2.5408	.082

To illustrate how significant the pre-and post-test scores were, histograms were used. The significant improvement in test scores from pre-test to post-tests is presented in Figure 1. As illustrated in the figure, the pre-test scores, with a mean of 4.84 and a standard deviation of 1.911, showed a relatively broad distribution, with most students scoring between 4 and 6 points. In contrast, the post-test scores had a higher mean of 9.76 and a slightly higher standard deviation of 2.5408, indicating higher average performance and more consistent scores. Notably, the most popular score range was from 7 to 12 points, demonstrating a marked shift towards higher performance across the 38 students.

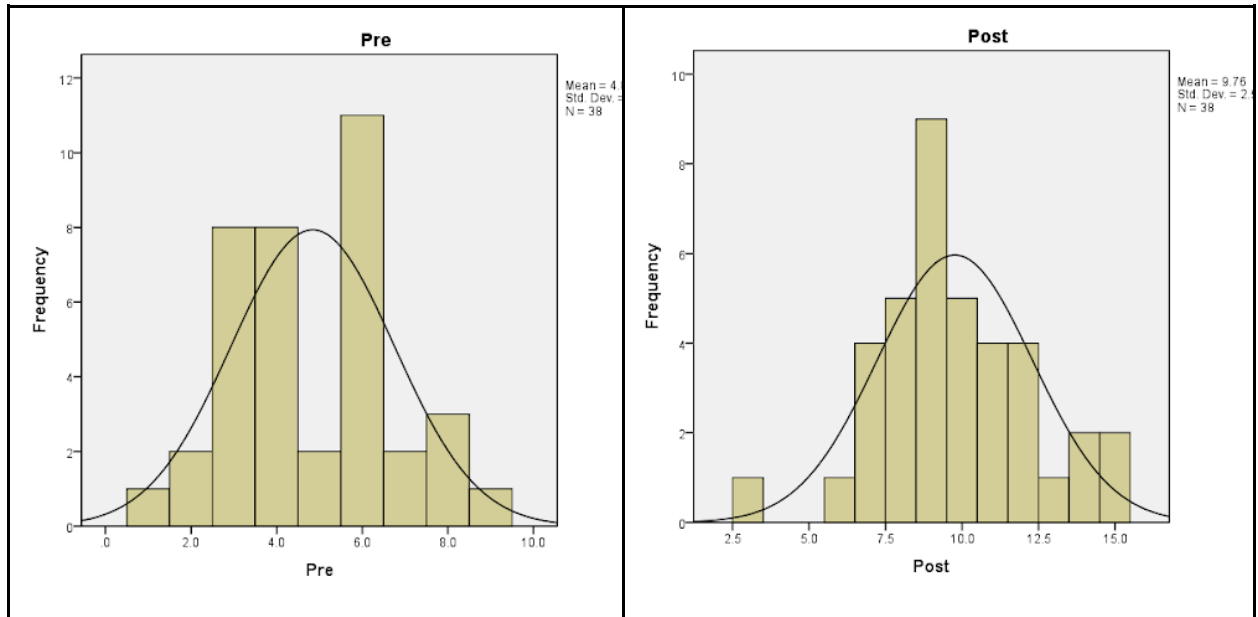


Figure 1: Frequency statistics for pre and post-tests

A paired samples t-test was conducted afterwards to assess the impact of the invention on improving writing skills, comparing pre and post-test scores after individual adoption of *Write & Improve*. The paired sample method indicates that, at baseline (M = 4.842, SD = 1.9105), the mean writing scores were subsequently enhanced after the intervention (M = 9.763, SD = 2.5408). These paired sample correlations depict a moderate and positive correlation between the pre and post-test scores ($r = .521, p > .001$), underscoring the consistency in writing performance.

Table 4: Comparison of mean scores of pre-test and post-test scores

	N	Mean	SD	t-Value	df	Correlation	Sig
Pre-test	38	4.842	1.9105	-13.502	37	.521	.001
Post-test	38	9.763	2.5408				

*Significant at $\alpha=0.05$

Table 5: Comparison of mean scores of pre-test and post-test employing sample test-paired differences

	Mean	SD	SD Error Mean	95% Confidence Interval of Difference		Sig. 2-tailed
				Lower	Upper	
PreTest - PostTest	-4.9211	2.2467	.3645	-5.6595	-5.6595	.000

*Significant at $\alpha=0.05$

As indicated in Tables 4 & 5 above, the paired samples test results highlight a statistically significant mean difference in writing scores between pre-test and post-test (M = -4.9211, SD = 2.2467, 95% CI [-5.6595, -5.6595], $p = .000$). These findings provide robust evidence of a significant improvement in writing scores following the intervention, showing the effectiveness of using the application.

4.2 Students' perception towards the utilization of the *Write & Improve* application**4.2.1 Using *Write & Improve* as a feedback tool**Table 6: Perceived use of *Write & Improve* as a feedback tool

No.	Statements	SD	D	N	A	SA	Mean	Cronbach Alpha
1	This app provides quick feedback.	2.56%	2.56%	10.26%	43.59%	38.46%	4.16	.935
2	This app provides accurate feedback.	2.56%	2.56%	12.82%	58.97%	20.51%	3.95	
3	This app provides detailed feedback.	2.56%	5.13%	20.51%	53.85%	15.38%	3.76	
4	This app provides more detailed feedback than the classroom teacher.	2.56%	17.95%	23.08%	38.46%	15.38%	3.47	
5	This app provides easy-to-understand feedback.	2.56%	5.13%	28.21%	43.59%	17.95%	3.71	

SD= Strongly Disagree; D= Disagree, N= Neutral; A= Agree; SA= Strongly Agree

The data collected from Table 6 indicates that a considerable majority of participants either agreed or strongly agreed with the statements concerning the feedback provided by the tool. For example, when questioned about the immediate feedback pace, over 80% of participants agreed with the claim that the app can provide timely feedback. Similarly, regarding feedback precision, 79.48% agreed or strongly agreed that the app provides accurate feedback, followed by 69.23% of participants who either agreed or strongly agreed that the app could generate comprehensive feedback. Additionally, regarding the comparison between the app's detailed feedback and classroom teachers, a significant proportion of 53.84% either agreed or strongly agreed that the tool offers more detailed feedback, with more than 50% of the students indicating their agreement. Finally, regarding the clarity of the feedback, more than 60% of the participants believed that the feedback given by this tool was straightforward; however, a small percentage of the students had opposite viewpoints. These results consistently highlight positive user perceptions regarding the quality and user-friendliness of the tool.

4.2.2 Perception of the usefulness of using the *Write & Improve* application to improve writing skillsTable 7: Overall perception of using *Write & Improve* in practising writing skills

No.	Statements	SD	D	N	A	SA	Mean	Cronbach Alpha
1	I enjoy practising writing more.	7.69%	2.56	33.33%	48.72	5.13%	3.42	.940
2	I can see that the quality of my writing has improved.	5.13%	2.56%	17.95%	53.85%	17.95%	3.79	
3	I feel that I can practice more after school.	5.13%	2.56%	12.82%	56.41%	20.51%	3.87	
4	I like this tool.	7.69%	0.00%	12.82%	66.67%	10.26%	3.74	
5	I find this helpful tool.	7.69%	0.00%	12.82%	51.28%	25.64%	3.89	
6	I find this tool easy to use.	5.13%	0.00%	20.51%	43.59%	28.21%	3.92	
7	I notice that the number of mistakes I make has decreased.	5.13%	5.13%	20.51%	48.72%	17.95%	3.71	
8	I can see the level of my writing thanks to the score given by <i>Write and Improve</i> .	5.13%	2.56%	15.38%	48.72%	25.64%	3.89	

SD= Strongly Disagree; D= Disagree, N= Neutral; A= Agree; SA= Strongly Agree

Table 7 shows that the students had varied responses regarding their experience with Write & Improve, a writing practice tool. To be more specific, approximately 75% of the participants noted an improvement in the quality of their writing. A similar proportion preferred the tool, indicating a favourable reception. Moreover, about two-thirds of the students found the tool easy to use. Similarly, the same number of them acknowledged a decrease in mistakes and agreed that they could discern the level of their writing thanks to the score given by *Write & Improve*. However, only around half of the students reported enjoying using this tool for additional practice sections.

4.2.3 Challenges of using Write & Improve in practicing writing skills

The participants had a range of viewpoints when applying the feedback system of the tool. In essence, a minority, representing 10.26%, expressed confusion over the feedback provided by the tool, with an additional 25.64% sharing similar uncertainties. Conversely, 25.64% reported understanding the comments, while 30.77% adopted a neutral point of view. Concerning the perception of feedback being generic and repetitive, another minority of 10.26% rejected this opinion, with 12.82% expressing similar feelings. Meanwhile, 23.08% remained neutral, while 43.59% agreed with this perception.

Additionally, 7.69% believed that most feedback addressed vocabulary and grammar, overlooking ideas, arguments, or writing organization.

In contrast, 51.28% approved this viewpoint, with an additional 7.69% strongly supporting it. A minority, 17.95%, faced challenges using the tool due to the absence of a laptop, while 25.64% shared this perspective.

On the other hand, 20.51% maintained a neutral standpoint, and 10.26% disagreed with this statement. Lastly, 20.51% found specific feedback inaccurate or mismatched with their intended expression, while 38.46% agreed. Comparatively, only 31% claimed that the app tended to give false feedback or misleading content. The data was then summarized in the table below:

Table 7: Challenges faced during use

No.	Statements	SD	D	N	A	SA	Mean	Cronbach Alpha
1	I don't understand the feedback given by the tool.	10.26%	25.64%	30.77%	25.64%	5.13%	2.89	.910
2	I understand the feedback but don't know how to revise my writing.	7.69%	10.26%	20.51%	46.15%	12.82%	3.47	
3	The feedback is general and repetitive.	10.26%	12.82%	23.08%	43.59%	7.69%	3.26	
4	Most feedback only corrects vocabulary and grammar without commenting on ideas, arguments, or the organization.	7.69%	2.56%	28.21%	51.28%	7.69%	3.50	
5	I have difficulty using the tool because I don't have a laptop.	17.95%	25.64%	23.08%	20.51%	10.26%	2.79	
6	Some feedback may be inaccurate or not aligned with what I want to express.	7.69%	20.51%	38.46%	23.08%	7.69%	3.03	

SD= Strongly Disagree; D= Disagree, N= Neutral; A= Agree; SA= Strongly Agree

4.2.4 Students' overall experience towards the application of Write & Improve in L2 writing classrooms

Most students found *Write & Improve* to be highly beneficial for enhancing their writing abilities. They appreciated the ease of use and the immediate feedback the tool provided on grammatical errors. As one

student commented, "*I find this application very interesting and convenient. I will use it when I study English,*" or "*The application helps us improve our writing skills significantly by checking and scoring our essays. It can also point out grammatical errors,*" as another participant affirmed.

Regarding the ability to provide detailed feedback and score essays, one student shared, "*This application is good because it allows us to practice writing and improve our skills. It points out grammatical mistakes so we can correct them.*" The app's potential to track students' progress and offer suggestions for better writing was also appreciated, as one user noted.

Write & Improve was praised as a practical tool for beginners and advanced learners, offering a user-friendly feature and valuable insights into writing skills. One student claimed, "*This is a handy tool for improving my writing skills, and I believe everyone should try it. After using the app under my teacher's guidance, I felt my writing ability improved.*" Several students expressed their interest in continuing to use the app.

Students identified some challenges while using *Write & Improve* despite the positive feedback. One common issue was the difficulty in understanding some of the feedback provided by the app. One student pointed out, "*Sometimes the feedback is hard to understand, making it difficult to figure out the corrections.*" Another user expressed dissatisfaction, stating, "*I don't find this app very useful; it is quite complicated.*" Additionally, some students felt that the features were limited in certain areas, with one user mentioning, "*Although it helps with spelling mistakes, it doesn't improve my writing style.*" However, teachers play the critical role in making things easier for students. As one student mentioned, "*Some new features can be too difficult at first, but with the teacher's guidance, I gradually adapted.*"

5 DISCUSSION

In terms of the impact of *Write & Improve* on students' writing scores, data from pre- and post-tests revealed a significant increase in mean writing scores from 4.842 to 9.763, as measured on a scale from 1 to 15 (Refer to Appendix A for the Cambridge Assessment Scale used for the pre- and post-tests assessment), indicating a substantial improvement in writing skills. This aligns with positive outcomes observed in related research by Abalkader (2022), Karpova (2020), Li et al. (2015), where students demonstrated improvements following the intervention using the tool, showing the effectiveness of applying technological tools in writing skill development.

Students' responses varied across behavioural and emotional perspectives regarding the perceived advantages and disadvantages of *Write & Improve* implementation. As shown in Table 5, promptness and accuracy emerged as the two prominent factors, with many students in this study acknowledging these aspects, followed by detailed and easy-to-understand feedback provided by the tool. These results corroborate with studies by Zhu et al. (2020), Ranalli et al. (2017), Li, Link, and Hegelheimer (2015), Lim & Phua (2019), which emphasized the utility of digital tools in offering instant feedback and stimulating idea generation. Most students expressed a desire to use the tool further due to its user-friendly nature and helpfulness, with half indicating increased enjoyment in practising writing skills and a decline in errors. Overall, students had positive perceptions regarding the significance of *Write & Improve* for their studies and their future use. These results were consistent with findings by Abalkader (2022) and O'Neill and Russell (2019) regarding positive feedback on comments and corrections received during use.

Despite the benefits, concerns arose among students regarding utilising *Write & Improve*. As seen in Table 7, more than half of the students reported difficulty revising their writing as the app did not provide detailed feedback, primarily on grammar and vocabulary corrections. Additionally, students noted that the language used in the system was overly general and repetitive. These findings shared some similarities with those collected by Miranty et al. (2021), highlighting the ineffective learning style observed when students lacked teacher instructions and guidance.

6 CONCLUSION & LIMITATION

The current study employed a mixed-methods approach to explore students' performance in writing skills and their attitudes towards the AWE tool used during a General English 1 course. The findings indicated a significant improvement in post-test scores compared to pre-test scores. Besides, the survey on perceptions suggested that the non-English majors generally had a favourable view of this AWE tool. The students

reported positive perceptions of ease of use and intended to continue using this tool to personalize their learning.

Although several challenges were noted, the findings indicated that learners support integrating this tool into the language teaching and learning environment. A positive attitude towards the tool is a stepping stone for language learners to accept and use it frequently in their learning journey, which is beneficial and crucial for language acquisition.

The overall findings of this study suggest that integrating this tool into language teaching and learning is advantageous for both students and teachers. The tool saves time for teachers by providing timely and regular feedback on students' writing tasks. For students, it allows them to practice their writing skills independently with prompted input and suggestions for revision. By using this tool, students can personalize their learning process.

This study has some limitations. Firstly, the number of participants was limited to thirty-eight. A larger sample size would be more beneficial in elucidating the impact of the tool on the proficiency of L2 students. Secondly, the experimental design was conducted by one of the researchers. Involving other teachers in the study could minimize possible biases and ensure methodological rigour for the entire study. Another limitation pertains to constraints related to the participants' schedules. Many students had a packed timetable, which might have affected their interest level and commitment to practising with the AWE tool. One last point is that further research should investigate L2 learners' task engagement and individual regulation when using this AI tool.

REFERENCES

- [1] M. S. Ariyanto, N. Mukminatien, and S. Tresnadewi. College students' perceptions of an automated writing evaluation as a supplementary feedback tool in a writing class, *Jurnal Ilmu Pendidikan*, vol. 27, no. 1, p. 41, 2021. Doi: 10.17977/um048v27i1p41-51.
- [2] T. Barker and F. Kemp. Network theory: A postmodern pedagogy for the written classroom, in *Computers and community: Teaching composition in the twenty-first century*, C. Handa, Ed. Portsmouth, NH: Heinemann, 1990, pp. 1-27.
- [3] P. Black and D. Wiliam. Assessment and Classroom Learning, *Assessment in Education: Principles, Policy & Practice*, vol. 5, no. 1, p. 7, 1998.
- [4] D. Chun, Using computer networking to facilitate the acquisition of interactive competence, *System*, vol. 22, no. 1, pp. 17-31, 1994.
- [5] K. Karpova. Integration of "Write and Improve" AWE Tool into EFL at Higher Educational Establishment: Case Study, *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, vol. 7, no. 2, pp. 137-150, 2020.
- [6] M. D. Langston and T. Batson. The social shifts invited by working collaboratively on computer networks: The ENFI project, in *Computers and community: Teaching composition in the twenty-first century*, C. Handa, Ed. Portsmouth, NH: Boynton/Cook, 1990, pp. 149-159.
- [7] F. V. Lim and J. Phua. Teaching writing with language feedback technology, *Computers and Composition*, vol. 54, 2019, Art. no. 102518.
- [8] P. L. Muliani and W. Widana. *Uji Persyaratan Analisis: Analisis Standar Pelayanan Minimal Pada Instalasi Rawat Jalan Di RSUD Kota Semarang*, 2020.
- [9] R. O'Neill and A. Russell. Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly, *Australasian Journal of Educational Technology*, vol. 35, no. 1, 2019.
- [10] M. C. Pennington. Computer-assisted writing on a principled basis: The case against computer-assisted text analysis for non-proficient writers, *Language and Education*, vol. 7, no. 1, pp. 43-59, 1993.
- [11] J. K. Peyton. Technological innovation meets institution: Birth of creativity or murder of a great idea? *Computers and Composition*, vol. 7, Special Issue, pp. 15-32, 1990.
- [12] R. Sidman-Taveau and M. Milner-Bolotin. Constructivist Inspiration: A Project-Based Model for L2 Learning in Virtual Worlds, *Texas Papers in Foreign Language Education*, vol. 6, no. 1, pp. 63-82, 2001.
- [13] N. Sullivan and E. Pratt. A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom, *System*, vol. 24, no. 4, pp. 491-501, 1996.
- [14] M. Warschauer. Comparing face-to-face and electronic communication in the second language classroom, *CALICO Journal*, vol. 13, no. 2, pp. 7-26, 1996.

- [15] M. Warschauer. Online learning in second language classrooms: An ethnographic study, in *Network-based language teaching: Concepts and practice*, M. Warschauer and R. Kern, Eds. New York: Cambridge University Press, 2000.
- [16] M. Warschauer. Technology and writing, in *The International Handbook of English Language Teaching*, C. Davison and J. Cummins, Eds. Norwell, MA: Springer, 2007, pp. 907-912.
- [17] Y. M. Wang. E-mail dialogue journaling in an ESL reading and writing classroom, Ph.D. dissertation, University of Oregon, Eugene, 1993.
- [18] *Write & improve with Cambridge*. (n.d.). Write & Improve with Cambridge English. Accessed: May 14, 2024. [Online]. Available: <https://writeandimprove.com/>

NÂNG CAO KỸ NĂNG VIẾT VÀ THÁI ĐỘ CỦA SINH VIÊN THÔNG QUA SỬ DỤNG CÔNG NGHỆ NHẬN XÉT TỰ ĐỘNG (AUTOMATED WRITING EVALUATION): NGHIÊN CỨU TẠI MỘT TRƯỜNG ĐẠI HỌC Ở VIỆT NAM

LE THI THIEN PHUOC¹, NGUYEN LAM ANH DUONG¹, PHAM THI THU¹, DANG HOANG MAI¹,
NGUYEN XUAN HONG*²

¹ Khoa Ngoại ngữ - Trường Đại học Công nghiệp TP HCM

² Trường Đại học Công nghiệp TP HCM

*Tác giả liên hệ: nguyenxuanhong@iuh.edu.vn

Tóm tắt. Nghiên cứu này nhằm tìm hiểu tính hiệu quả của công cụ đánh giá bài viết tự động (AWE), có tên là Write & Improve, trong việc nâng cao kỹ năng viết của sinh viên không chuyên ngành Ngôn ngữ Anh tại một trường đại học công lập ở Việt Nam, đồng thời đánh giá mức độ nhận thức của sinh viên về ưu điểm và nhược điểm của công cụ này. Đối tượng tham gia nghiên cứu bao gồm 38 sinh viên không chuyên ngành Ngôn ngữ Anh đang học môn Tiếng Anh 1. Nghiên cứu này dùng phương pháp nghiên cứu hỗn hợp, bao gồm việc thu thập dữ liệu thông qua các bài kiểm tra trước và sau khi sử dụng công cụ, bảng câu hỏi về nhận thức của sinh viên về công cụ Write & Improve sau một học kỳ sử dụng nó cũng như ý kiến chung của các bạn sinh viên về công cụ này. Những phát hiện của bài nghiên cứu đã chứng minh tính hiệu quả của công cụ này trong việc cải thiện kỹ năng viết của sinh viên và nêu bật nhận định tích cực của các em đối với việc sử dụng công cụ AWE trong các lớp học kỹ năng viết tiếng Anh. Bên cạnh đó, các dữ liệu cũng chỉ ra rằng sinh viên vẫn gặp phải một số thách thức. Chính vì vậy, giáo viên cần chú trọng đến những vấn đề sinh viên gặp phải khi tích hợp công cụ này vào quá trình rèn luyện kỹ năng viết cho sinh viên. Một số khuyến nghị và ý nghĩa thực tiễn cho việc dạy, học và đánh giá ngôn ngữ sẽ được thảo luận cụ thể trong bài báo này.

Từ khóa. Trí tuệ nhân tạo, nhận xét tự động, AWE, Viết và cải thiện.

Ngày nhận bài: 22/6/2024

Ngày chấp nhận đăng: 30/9/2024