Automated Writing Evaluation: Users' Perception and Expectations

Azman Che Mat*, Luqmanul Hakim Zulkornain, and Nurul Ajleea Abdul Rahman

UiTM Cawangan Terengganu, Dungun Campus, 23000 Dungun, Terengganu, Malaysia
Email: azman531@uitm.edu.my (A.C.M.); luqman562@uitm.edu.my (L.H.Z.); ajleaa@uitm.edu.my (N.A.A.R.)
*Corresponding author
Manuscript received August 3, 2023; revised September 1, 2023; accepted October 8, 2023; published February 2, 2024

***Abstract*—Automated Writing Evaluation (AWE) is a great tool for writing evaluation and correction due to its ability to analyze writing. Recent debates in writing have focused on the distinctions between the traditional method of teaching writing and the implementation of AWE in the teaching of writing and choosing the better of the two. The objective of this study is to explore the learners' perception of AWE among university students as well as their expectation of AWE implementation. The design of the study is descriptive in nature, employing the quantitative method with 72 university students participating in the experiment. This group underwent an intervention of 10 weeks being exposed to the AWE use. In this study, a questionnaire was developed as a tool for data collection. The implementation of AWE was viewed favourably by respondents. It was discovered that respondents consider AWE to be successful in terms of improving vocabulary use, language use, spelling, and punctuation. Respondents however, consider AWE to be less helpful in improving the substance and organisation of students' ESL writing. The findings indicate that using technology in this digital learning environment cannot sufficiently support the process of ESL writing assessment and still necessitates traditional evaluation methods such as educators explaining and helping their students organize their writing and develop successful material in ESL writing.**

***Keywords*—automated writing evaluation, digital learning, grammarly, technology, writing skills**

I. Introduction

Improving English language proficiency is of utmost importance in today's daily life. Among the four essential language skills—speaking, listening, reading, and writing—writing often presents the greatest challenge. Particularly for foreign students learning English, expressing ideas and emotions in a suitable manner can be daunting [1]. In traditional classroom settings, instructors play a crucial role in identifying and rectifying writing errors. Students rely on their instructors' guidance to address any flaws in their written work. Similarly, when striving to enhance their writing skills, students can greatly benefit from the valuable advice and support provided by their instructors, which contributes to their overall learning experience [2].

Writing is an inherently challenging task that requires meticulous attention and a deep understanding of the language in which it is being written. Maharani supports this notion by highlighting the difficulty associated with writing as a skill [3]. Faller further emphasizes this claim, noting that writing demands substantial cognitive and linguistic abilities [4] in which Ghonsooly and Shalchy also noted that cultural intelligence also influence writing ability [5]. Another notable point regarding language proficiency and writing skills, as discussed by Saglam and Duman [6], is that language competency played a vital role in effectively integrating source−based information into essays in writings which are source−based. Writing self−efficacy was also found to be associated with writing performance, the student's motivations to write, as well as their self−regulated learning [7].

While writing can be challenging even for proficient language users, it poses an even greater difficulty for students in a setting where English is taught as a second language. Yang acknowledges the obstacles faced by ESL students, including their linguistic limitations, inadequate language acquisition methods, and a lack of awareness regarding proper writing techniques [8]. These factors significantly impede their ability to produce well−crafted written pieces.

Consequently, achieving mastery in writing necessitates a close collaboration between students and teachers, as they work towards the common objective of producing well−crafted written pieces using appropriate writing techniques. Effective teamwork in this regard is not easily achieved; it requires dedicated time and effort. The process itself is multifaceted and varies from one student to another [9]. ESL students encounter challenges related to genre, structure, and meaning in their writing. Among these challenges, spelling errors are the most prevalent and often overlooked as mere accidents [10]. However, these spelling difficulties lead to another issue as they directly impact the intended meaning, resulting in written work that fails to effectively convey the intended message [11].

However, with the emergence of Education 5.0, the integration of the Internet of Things (IoT) has brought forth a new dimension to students' learning experiences. A vast number of students worldwide now have access to the internet as a platform for learning and acquiring knowledge. This has become even more crucial in light of the global pandemic, which necessitated the rapid adoption of online educational practices to ensure continuity in learning across the globe. The past two years have witnessed a significant shift towards utilizing the internet as a means to sustain educational activities worldwide [12].

One significant aspect of IoT's impact on education is the transformation in the way learners receive feedback on their learning progress. In particular, when it comes to writing, IoT has facilitated the integration of Automated Writing Evaluation (AWE) systems into the learning process. AWE leverages AI technologies to automatically analyze and evaluate written texts, providing learners with valuable feedback on their writing skills.

Automated Writing Evaluation (AWE) is an emerging application that offers multiple functions for detecting writing texts, including word selection, suggested phrases, grammar correction, and style analysis [13]. A significant aspect of

AWE is its ability to check text structures in English and provide alternative styles, allowing users to accept or reject the recommendations. AWE software is highly beneficial for both educators and students as it facilitates the analysis of ESL writing in teaching and learning contexts [14–16].

One of the most significant contributions of AWE software is its provision of formative feedback and evaluation. Modern AWE software includes components such as plagiarism checkers that support extensive instructional practice, which can greatly assist ESL students with their writing [17]. ESL learners often encounter challenges in spelling, punctuation, and grammar due to linguistic and educational constraints, leading to unintentional errors that go unnoticed [10, 18]. To address these challenges, incorporating grammar and spelling checks into English teaching and learning, particularly in writing instruction, is essential [19].

Grammarly, an exceptional tool, can be effectively utilized in ESL writing classes to facilitate grammar, spelling, punctuation, and plagiarism checks [20]. It serves as proofreading software that can detect grammatical errors, identify plagiarism, and offer suggestions for grammar, punctuation, and synonyms [21]. Previous research has examined the use of Grammarly, as demonstrated by Ghufron and Rosyida [22]. The main goal of this study is to investigate the feasibility of using language programming and appropriate vocabulary to provide educators with remedial guidance to reduce students' errors in ESL writing. Additionally, it aims to explore the impact of Grammarly's presentation on improving the quality of students' English writing, including its ability to enhance students' confidence by delivering automatic grammar feedback and offering alternative terms to improve their writing skills [23].

Previous research has also demonstrated the application of Grammarly in various writing genres, such as narrative texts, abstracts, and free writing, to help students enhance their writing skills [24–26]. These studies have highlighted Grammarly's effectiveness in recognizing student errors, boosting vocabulary development, and correcting punctuation [27].

Monitoring the writing process and providing students with relevant and helpful commentary is a time−consuming and subjective task [9]. To address this challenge, Grammarly and other computer−based programs are increasingly being utilized to assist individuals in improving their writing skills. The emergence of AI−powered writing tools, accessible on mobile devices, has the potential to help students acquire and develop writing skills that are often difficult to master through traditional education. Writing is a complex and emotional process, comparable to a distinguished scientific career [28].

Furthermore, English second language (ESL) students face additional challenges due to linguistic and educational constraints [29]. Unfortunately, higher education often fails to provide sufficient preparation for graduate students in this regard. Grammarly, as one of the tools available, automatically detects and corrects language structure, spelling, accentuation, and other writing errors. Despite the usefulness of AWE software, little research has been conducted on its effects on improving writing skills, and students' perceptions of AWE remain understudied. Therefore, it would be fascinating to conduct a study evaluating the efficacy of AWE among university students in developing writing abilities to gain a better understanding of automated writing evaluation, particularly Grammarly.

In the digital age, students' writing skills are evolving to a level where they must attain a higher proficiency to remain competitive in the global market. Traditional language instruction faces the challenge of shifting towards approaches that extend beyond the limitations of traditional classroom settings. In the era of IoT, it is important to note the prevalence of technology such as smartphones and the evolving landscape of technology trends such as new media usage, such as Google, gaming, Instagram, and YouTube to accommodate the new generation of students as it could also help in predicting these digital natives' foreign language learning effort [30, 31]. Innovative materials, such as Automated Writing Evaluation (AWE), can be effectively employed for teaching and learning purposes as it helps with self-assessment which empowers students with autonomy in writing [32] thus boosting their confidence to write.

However, it is important to note that AWE software, notably recognized during the pandemic [33], cannot replace the guidance provided by educators, as students still require their support to improve the content of their work [34]. This is important as AWE is still developing that it has not achieved the full potential to provide the best evaluation on the structure of the text, logic, as well as the coherence in the writing [35]. Since AWE is dependent on algorithm, students writing for test will tend to put more emphasis on the strategies to beat the algorithm to pass the test. This beats the purpose of learning writing itself which is based on the knowledge and skills [36]. Since there are many shortcomings in the use of AWE even though AWE does help in facilitating the writing process as well as the analysis of the writing, the best way is to integrate AWE into the traditional writing teaching and process.

While the primary goal of AWE is to aid the writing process in general, it is crucial to recognize its potential to support the learning process and enhance students' overall writing skills.

This study aims at achieving two objectives as follows:

RO1: To identify student perception of writing self−proficiency, before and after intervention session. This objective has developed a hypothesis:

H0: There is no significant difference in respondents' perceptions of their writing proficiency before and after the intervention

H1: There is a significant difference in respondents' perceptions of their writing proficiency before and after the intervention

RO2: To explicitly understand student's expectation of the use AWE in their writing learning process. This objective has developed a hypothesis as:

H0: There is no significant difference in respondents' expectation of AWE assistance in their writing learning before and after the intervention

H1: There is a significant difference in respondents' perceptions of AWE assistance in their writing learning before and after the intervention.

II. METHODOLOGY

A. Research Design

Based on quantitative approach, this study has a descriptive nature and employs statistical analysis of the data to examine the effects of an intervention. This method was employed to statistically analyze the data gathered from the population sample to find the common patterns. Since the study only looks into learners' perception of AWE among university students as well as their expectation of AWE implementation, the quantitative method was employed as it is deemed sufficient for such application as the study did not really dive too deep into why such phenomenon occured. Specifically, a pre−post experiment was conducted on a single group of participants. The sampling method used purposive sampling method where the participants we're selected purposely based on the criteria aligned with the research objectives which are to find learners perception and expectation of AWE. The intervention involved exposing this group to the use of AWE for a period of 10 weeks.

B. Participants

During the implementation of online classrooms, obtaining a suitable sample for the study posed a challenge. The shift to virtual learning has impacted students' willingness to engage in additional assignments. Nevertheless, the study was fortunate to have a number of students who were willing to participate. The study includes a treatment group that was assigned to complete a questionnaire before and after a specific period. The Table 1 presents the number of participants involved in the study.

Table 1. Participants distribution based on universities

Institution	Participant Total (N)
UiTM	32
UMK	15
UMP	11
UM	14
Total (N)	72

The study involved a total of 72 participants from four universities in Malaysia: Universiti Teknologi Mara Cawangan Terengganu (UiTMCTKD), Universiti Malaya (UM), Universiti Malaysia Kelantan (UMK), and Universiti Malaysia Pahang (UMP). To ensure impartiality, the participants were purposely not selected from English language−related disciplines such as English language studies, English literature, English for communication, and Teaching English as a Second Language (TESL). The objective was to assess the perception and expectations of AWE, specifically Grammarly, in a general context. All participants were enrolled in writing classes at the time of the study.

C. Instruments

In this study, questionnaire set was developed as a data collection tool. Conventionally, the questionnaire set consisted of demographic profiles of respondents and variables of the study. As mentioned earlier, participants of the study were an experimental group. Therefore, this study has developed a questionnaire adapted from [26] and [32] according to the two objectives of the study as demonstrated by the Table 2 below:

Table 2. Items of questionnaire for RO1 and RO2

No.	Item
RO1	**PERCEPTION ON SELF PROFICIENCY IN ENGLISH LANGUAGE (WRITING)**
1.	I can use appropriate vocabulary and word forms to effectively communicate with the reader.
2.	I can use appropriate spelling, capitalization, and punctuation.
3.	I can write an accurate summary of information that I have read in English.
4.	I can write a good academic writing.
5.	I can write a good conclusion for an English essay.
6.	I can effectively brainstorm to gather ideas before writing.
7.	I can write quickly in English.
8.	I write for pleasure in my free time in English.
9.	I reward myself when I have finished writing.
10.	I make notes or try to remember feedbacks I get on my writing so I can use it the next time I write.
RO2	**PERCEPTION ON AWE (BENEFITS OF USING AWE IN LEARNING PROCESS)**
1.	Improve spelling
2.	Improve grammar
3.	Improve sentence structure
4.	Improve punctuation
5.	Improve vocabulary
6.	Improve writing style
7.	AWE gives detailed feedback.
8.	AWE makes helpful suggestions for improving my work.
9.	AWE gives good explanations about my errors.
10.	AWE helps me understand grammar rules.
11.	I do agree that students get usefulness with the use of AWE in writing class.
12.	I do agree that students get usefulness with the use of AWE in writing class
13.	AWE is a better choice for evaluating your writing compared to a human evaluator.
14.	I prefer AWE over human evaluator.

D. Data Collection

After identifying that the participants were willing to cooperate, a set of questionnaires was distributed to them for pretest. Participants were given ten weeks of exposure to and to use AWE. After the intervention, another set of questionnaires was distributed to look into participants' perception of AWE use for their writing.

The distribution of the questionnaire was carried out through Google form as it was the most convenient platform for both researcher and participants of the study. Participants were required to sincerely respond without any interference from others.

E. Data Analysis

The collected data were analyzed descriptively to answer the research questions. By using an Excel sheet and SPSS version 27, the calculation was manually carried out to alter the data statically resulting in percentage, mean scores and T−test analysis. Then each related finding was described with a concentration on the objectives of the study.

III. FINDINGS

The collected data will be analyzed descriptively to answer the research questions. By using an Excel sheet and SPSS version 27, the calculation was manually carried out to alter the data statically resulting in percentage, mean scores and T−test analysis. Then each related finding was described with

a concentration on the objectives of the study.

This section will elaborate the objectives of the study based on data analysis. To begin with, the demographic profile of the participants should be demonstrated to give clear evidence of the source of the data collected in this study. As mentioned in methodology, this study has obtained 72 participants from four universities, namely UiTM, UMP, UMK and UM. As for demographic profile, four criteria were identified as Table 3.

Table 3. Demographic profile of the participants

No	Profile	Category	Frequency	Percentage
1.	Gender	Male	29	40.28%
		Female	43	59.72%
2.	Age Group	17–20	30	41.67%
		20–23	36	50%
		23–26	6	8.33%
		others	0	0%
3.	Study Level	Undergraduate	72	100%
4.	Sijil Pelajaran Malaysia (SPM) Result	A	21	29.17%
		B	32	44.44%
		C	9	12.5%
		D	7	9.72%
		E	1	1.39%
		others	2	2.78%

Table 3 above shows the number of male students is 29 (40.28%) while female students are 43 (59.72%). The age number of the participants indicated 30 (41.67%) for the age group 17–20, 36 (50%) for the age group 20–23 and 6 (8.33%) for the age group 23–26. In this study, because participation depended on a voluntary basis, it was hard to obtain a balanced number of participants in different locations. Therefore a combination between diploma and degree students was made up to ensure a credible number of participants for the study. When it comes to academic background, especially for tertiary level in English subject, most of the participants have different grades which A is 21 (29.17%) following B is 32 (44.44%), C is 9 (12.5%), D is 7 (9.27%) and E is only 1 (1.39%). Lastly, 2 (2.78%) did not have taken SPM for their different background.

A. Research Objective 1: Students' Perception on Writing Self-Proficiency

The first objective of the study is to identify students' perception on writing self−proficiency. Ten items of the questionnaire represent the perception of the participants. Pre and post questionnaire responses were analyzed to scrutinize the level of their perception before and after the exposure to the use of Grammarly during the semester.

Table 4 presents the participants' enhanced perception of their writing abilities. Initially, all participants considered their writing to be satisfactory. However, upon exposure to Automated Writing Evaluation (AWE), certain aspects emerged that warranted closer examination. For instance, in item no. 2, participants noted that the use of AWE had improved their proficiency in appropriate spelling, capitalization, and punctuation, with the mean increasing from 3.805 to 4.111. for item no. 8, the mean for "writing for pleasure in free" time also increases after exposed to AWE from 2.986 to 3.583. In conclusion, the participants demonstrated a positive inclination towards the advantages of AWE, be it in terms of enhancing their writing skills or transforming their mindset with regards to writing in English.

Table 4. Paired samples statistics for student perception on writing self− proficiency

No.	Item	Test	Mean	N	Std. Dev	Std. Error Mean
1.	I can use appropriate vocabulary and word forms to effectively communicate with the reader.	PRE	3.8056	72	0.72460	0.08539
		POST	4.1111	72	0.70322	0.08288
2.	I can use appropriate spelling, capitalization, and punctuation.	PRE	3.8056	72	0.72460	0.08539
		POST	4.1111	72	0.70322	0.08288
3.	I can write an accurate summary of information that I have read in English.	PRE	3.5556	72	0.66901	0.07884
		POST	4.0000	72	0.65003	0.07661
4.	I can write a good academic research paper.	PRE	3.4444	72	0.66901	0.07884
		POST	3.7639	72	0.74101	0.08733
5.	I can write a good conclusion for an English essay.	PRE	3.5833	72	0.72675	0.08565
		POST	3.8750	72	0.69073	0.08140
6.	I can effectively brainstorm to gather ideas before writing.	PRE	3.5972	72	0.74417	0.08770
		POST	3.9444	72	0.70987	0.08366
7.	I can write quickly in English.	PRE	3.3611	72	0.81024	0.09549
		POST	3.7222	72	0.87568	0.10320
8.	I write for pleasure in my free time in English.	PRE	2.9861	72	0.88003	0.10371
		POST	3.5833	72	0.96049	0.11319
9.	I reward myself when I have finished writing.	PRE	3.3889	72	1.0947	0.12903
		POST	3.7917	72	0.99205	0.11691
10.	I make notes or try to remember feedbacks I get on my writing so I can use it the next time I write.	PRE	3.6389	72	0.99726	0.11753
		POST	3.8750	72	0.83813	0.09877

To test hypothesis one of this study, Paired T−test analysis was carried out to determine the hypothesis. Table 5 displays the result of the analysis.

Ten items were tested for the hypothesis for RO1 (To identify students' perception on writing self−proficiency before and after intervention session). These items can be broken down into 3 categories which are components of writing, ability to summarise and conclude, and lastly,

satisfaction in writing. These are demonstrated as follow:

1) Components of writing (Pair 1 and Pair 2)

The first category which includes Pair 1 and Pair 2 is related to lexis and its forms as well as its spelling, capitalization, and punctuation.

A paired−samples T−test was conducted to compare the use of appropriate vocabulary and word forms to effectively communicate with the reader before and after the intervention. The result illustrates that there was a significant difference in the comparison of before (M = 3.8, SD = 0.72) and after (M = 4.1, SD = 0.70) the intervention; the level of significance value is 0.05 (at the 0.05 level of significance, t(71) = −3.57, p = 0.001.

The second paired−samples T−test was conducted to compare the use of appropriate spelling, capitalization, and punctuation before and after the intervention. Here, comparison of the before and after exposure to AWE, the result also illustrates significant difference where before exposure to AWE, M = 3.8 and SD = 0.72, and after exposure to AWE, M = 4.1 and SD = 0.70 while the intervention, t(71) = −3.57 and p = 0.001.

The results from the T−test confirms the H1 for RO1 in which there is a significant difference in respondents' perceptions of students' writing proficiency before and after the intervention. These results suggest participants' perception on their writing becomes more positive after exposure to AWE.

2) Ideating, summarizing, and concluding a writing (Pair 3 to Pair 6)

The second category is on the ability to ideate, summarize, and conclude a writing. This includes Pair 3 to Pair 6.

For Pair 3, a T−test was conducted to compare the ability to write an accurate summary of information read in English before and after the intervention. Here, the difference is significant as can be observed that before the exposure to AWE, M = 3.6 and SD = 0.67 while after exposure, M=4.0 and SD=0.65 and the intervention t(71) = −5.31 and p = 0.000.

T−test for Pair 4 was conducted to compare the ability to write a academic research paper. Before exposure to AWE, M = 3.5 and SD = 0.66. After exposure to AWE, M = 3.8 and SD = 0.74. The intervention here, t(71) = −3.30 and p = 0.001, which show significant differences.

To compare the ability to write a good conclusion for an English essay before and after the intervention, T−test was conducted on Pair 5. A significant difference can be observed here where before exposure to AWE, M = 3.6 and SD = 0.73. After exposure, M = 3.9 and SD = 0.69, and the intervention, t(71) = −3.11 and p = 0.003.

A paired−samples T−test was conducted on Pair 6 to compare the ability to effectively brainstorm to gather ideas before writing essay before and after the intervention. Before exposure to AWE, M = 3.6 and SD = 0.73, and after exposure, M=3.9 and SD=0.69. The intervention, t(71) = −3.50 and p = 0.001. The results reveal that there is a significant difference.

The results from the T−test confirms the H1 for RO1 in which there is a significant difference in respondents' perceptions of students' writing proficiency before and after the intervention. These suggest that intervention does have an impact on participants. Specifically, the results suggest that when exposed to AWE, participants' perception on their writing becomes more positive.

3) Attitude pertaining to writing (Pair 7 to Pair 10)

The third category is on the attitude of the population towards writing. The category includes Pair 7 to Pair 10.

T−test was conducted on Pair 7 to compare the ability to write quickly in English before and after the intervention. Before exposure to AWE, M = 3.4 and SD = 0.81. After exposure, M = 3.7 and SD = 0.89. The intervention, t(71) = −3.50 and p = 0.001. Hence there is a significant difference.

For Pair 8, a T−test was conducted to compare the ability to write for pleasure in free time in English before and after the intervention. There is a significant difference where before exposure to AWE, M=3.0 and SD=0.88, and after exposure, M=3.6 and SD=0.96. The intervention is t(71) = −4.90 and p = 0.000.

The T−test conducted on Pair 9 was to compare the practice of self−rewarding when writing is finished, before and after the intervention. From the results before exposure (M = 3.4, SD = 1.09) and after exposure to AWE (M = 3.8, SD = 1.00) with the intervention, t(71) = −3.62, p = 0.001, it could be observed that there is in fact a significant difference.

The result of the paired−samples T−test conducted on Pair 10 is an interesting one as it is opposed to other 9 pairs. It was conducted to compare the practice of making notes or trying to remember feedbacks on a writing and using them in the next writing, before and after the intervention. There was no significant difference before exposure (M = 3.6, SD = 1.00) and after exposure to AWE (M = 3.9, SD = 0.10) as the intervention, t(71) = −1.98 and p = 0.052 (more than 0.05).

The results from the T−test confirms the H1 for RO1 in which there is a significant difference in respondents' perceptions of students' writing proficiency before and after the intervention. The results for most items, except for Item 10, are significant in which they are lower than 0.05. This depicts that students perceived that their writing skills have improved after the intervention.

B. Research Objective 2: Student's Expectation of the Use AWE in Their Writing Learning Process

The second objective of the study is to identify the respondents' expectation of the use of AWE for their writing. Before going directly to their response regarding the research objective, their exposure to the use of AWE previously were posted including experience using the AWE, knowledge about AWE and types of AWE they were familiar with.

Table 6 demonstrates the frequency of respondents for their experience which clearly stated that most of them are likely not getting used to AWE. The response from 72 respondents indicated only 34.72% or 25% agreed they have used the AWE. A total 36 out of 72 respondents admitted they do not know what AWE is. Furthermore, the table shows the most popular type of AWE known by respondents. By following the sequence from top to bottom, Grammarly was the most popular followed by Language Tool and SpellChecker. net, each with 27% and 22.22%. Then, rest of the lists of AWE commonly were not aware by the respondents which indicated by 0–12%.

Table 5. Paired samples test

Pair	Item	Paired Differences					T	df	Sig. (2−tailed)
		Mean	Std. Dev	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
1	I can use appropriate vocabulary and word forms to effectively communicate with the reader.	−0.30556	0.72460	0.08539	−0.47583	−0.13528	−3.578	71	0.001
2	I can use appropriate spelling, capitalization, and punctuation.	−0.30556	0.72460	0.08539	−0.47583	−0.13528	−3.578	71	0.001
3	I can write an accurate summary of information that I have read in English.	−0.44444	0.70987	0.08366	−0.61126	−0.27763	−5.313	71	0.000
4	I can write a good academic research paper.	−0.31944	0.81925	0.09655	−0.51196	−0.12693	−3.309	71	0.001
5	I can write a good conclusion for an English essay.	−0.29167	0.79501	0.09369	−0.47849	−0.10485	−3.113	71	0.003
6	I can effectively brainstorm to gather ideas before writing.	−0.34722	0.84186	0.09921	−0.54505	−0.14940	−3.500	71	0.001
7	I can write quickly in English.	−0.36111	0.87702	0.10336	−0.56720	−0.15502	−3.494	71	0.001
8	I write for pleasure in my free time in English.	−0.59722	1.04355	0.12298	−0.84245	−0.35200	−4.856	71	0.000
9	I reward myself when I have finished writing.	−0.40278	0.94436	0.11129	−0.62469	−0.18086	−3.619	71	0.001
10	I make notes or try to remember feedbacks I get on my writing so I can use it the next time I write.	−0.23611	1.01389	0.11949	−0.47436	0.00214	−1.976	71	0.052

Table 6. Respondents profile with AWE

No	Item	Respond	Percentage
1.	Have you ever used AWE?	Yes	25(34.72%)
		Maybe	16(22.22%)
		No	31(43.06%)
2.	Do you know what AWE is?	Yes	19 (26.34%)
		Maybe	17 (23.61%)
		No	36 (50%)
3.	Types of AWE you know	Grammarly	70(97.22%)
		Criterion® Online Writing Evaluation Service	0(0%)
		MY Access!	2(2.78%)
		RightWriter	2(2.78%)
		WritePlacer	2(2.78%)
		Ginger Software	4(5.56%)
		Grammark	6(8.33%)
		Grammarcheckme	9(12.5%)
		LanguageTool	20(27.78%)
		PaperRater	0(0%)
		Queequeg	0(0%)
		Spellchecker.net	16(22.22%)
		SpellCheckPlus	6(8.33%)
		WhiteSmoke	0(0%)

In terms of improving writing, the findings in Table 7 show that every aspect of writing skills have improved after respondents were exposed to the use of AWE. Before the exposure, respondents believed that AWE can assist their writing whether in spelling 48(66.67%), grammar 66(91.67%), sentence structure 46(63.89%), punctuation 46(63.89%), vocabulary 43(59.72%) and even writing style 39(54.17%).

After getting exposure and having experience with AWE within 10 months, respondents have slightly increased their belief that AWE will assist them in writing such as in spelling 62(86.11%), grammar 71(98.61%), sentence structure 46(63.89%), punctuation 42(58.33%), vocabulary 47(65.28%) and even writing style 41(65.28%). Interestingly, the highest aspect that is believed to improve is the grammar aspect in pre and post−test. Whereby, the lowest aspect is the improvement of punctuation with 31(43.06%) before the exposure however has increased after the exposure.

According to Table 8, items of the questionnaire required the respondents to give their feedback regarding what AWE offers to them and their perception on AWE on its usefulness and use in writing. Based on the findings, 4 questions were asked about what AWE offers them to assist in their writing and 4 more questions on their perception on AWE on its usefulness and use in writing. Each item indicated growing scores after the intervention session with all the scores hitting 4.00 and above although all good score of no less than 3.00 were indicated in the pretest.

Table 7. How can AWE improve writing

No.	Items	Pretest	Post−test
1.	Improve spelling	48(66.67%)	62(86.11%)
2.	Improve grammar	66(91.67%)	71(98.61%)
3.	Improve sentence structure	46(63.89%)	46(63.89%)
4.	Improve punctuation	31(43.06%)	42(58.33%)
5.	Improve vocabulary	43(59.72%)	47(65.28%)
6.	Improve writing style	39(54.17%)	41(56.94%)

As demonstrated in the Table 8, for item 1: AWE gives detailed feedback received 3.79 before and 4.25 after intervention; item 2: AWE makes helpful suggestions for improving my work received 4.00 and 4.39; item 3: AWE gives good explanations about my errors received 3.29 and

4.22; Item 4: AWE helps me understand grammar rules received both 3.89 and 4.33. Item 5: I do agree that students get usefulness with the use of AWE in writing class indicated 4.06 and 4.38; item 6: I do agree that students get usefulness with the use of AWE in writing class with score 4.07 and 4.40; item 7: AWE is a better choice for evaluating your writing compared to a human evaluator received 3.50 and 4.03, the last item is item 8: showed slight difference compared to item 1−6 which is "I prefer AWE over human evaluator" with score 3.36 and 3.74.

Table 8. Paired samples statistics for student perception on awe assistance

No.	Item	Test	Mean	N	Std. Deviation	Std. Error Mean
1.	AWE gives detailed feedback.	PRE	3.7917	72	0.62658	0.07384
		POST	4.2500	72	0.70711	0.08333
2.	AWE makes helpful suggestions for improving my work.	PRE	4.0000	72	0.60514	0.07132
		POST	4.3889	72	0.64032	0.07546
3.	AWE gives good explanations about my errors.	PRE	3.9167	72	0.74588	0.08790
		POST	4.2222	72	0.73585	0.08672
4.	AWE helps me understand grammar rules.	PRE	3.8889	72	0.68290	0.08048
		POST	4.3333	72	0.73158	0.08622
5.	I do agree that student get usefulness with the use of AWE in writing class.	PRE	4.0556	72	0.60255	0.07101
		POST	4.3750	72	0.65944	0.07772
6.	AWE is easy to use, especially in writing class.	PRE	4.0694	72	0.65706	0.07744
		POST	4.4028	72	0.66417	0.07827
7.	AWE is a better choice for evaluating your writing compared to human evaluator.	PRE	3.5000	72	0.80491	0.09486
		POST	4.0278	72	0.87165	0.10273
8.	I prefer AWE over human evaluator.	PRE	3.3611	72	0.99726	0.11753
		POST	3.7361	72	1.03452	0.12192

Table 9. Paired samples test

Pair	Item	Paired Differences					T	df	Sig. (2−tailed)
		Mean	Std. Dev	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
1	AWE gives detailed feedback.	−0.45833	0.80382	0.09473	−0.64722	−0.26944	−4.838	71	0.000
2	AWE makes helpful suggestions for improving my work.	−0.38889	0.79710	0.09394	−0.57620	−0.20158	−4.140	71	0.000
3	AWE gives good explanations about my errors.	−0.30556	0.91373	0.10768	−0.52027	−0.09084	−2.838	71	0.006
4	AWE helps me understand grammar rules.	−0.44444	0.74850	0.08821	−0.62033	−0.26856	−5.038	71	0.000
5	I do agree that student get usefulness with the use of AWE in writing class.	−0.31944	0.68846	0.08114	−0.48123	−0.15766	−3.937	71	0.000
6	AWE is easy to use, especially in writing class.	−0.33333	0.76912	0.09064	−0.51407	−0.15260	−3.677	71	0.000
7	AWE is a better choice for evaluating your writing compared to human evaluator.	−0.52778	0.93405	0.11008	−0.74727	−0.30829	−4.795	71	0.000
8	I prefer AWE over human evaluator.	−0.37500	1.04055	0.12263	−0.61952	−0.13048	−3.058	71	0.003

To test hypothesis two of this study, Paired T−test analysis was carried out to determine the hypothesis. Table 9 above displays the result of the analysis: hypothesis for RO2 (To explicitly understand student's expectation of the use AWE in their writing learning process). These items of the questionnaire were grouped into two themes; the first being what AWE offers to assist in writing and the second is perception on AWE on its usefulness and use in writing. The first theme comprises of Item 1 to Item 4 while the second theme are made up of Item 5 to Item 8. The results are demonstrated as the followings:

1) Theme 1—What AWE offers to assist in writing

A paired−samples T−test was conducted on Item 1 to compare the perception on the thoroughness of the feedbacks given by AWE before and after the intervention. Before intervention, M = 3.8 while SD = 0.72. After intervention, M = 4.1 and SD = 0.60. At the level of significance value is 0.05, t(71) = −4.80 and p = 0.000. Hence, it shows a great significance. These results suggest that intervention does exhibit an impact where after being exposed to the use of AWE, participant perceived that the feedbacks provided by AWE are more detailed.

The paired−samples T−test conducted on Item 2 compared the perception on the usefulness of the suggestion made by AWE before and after the intervention. The result indicated significant difference where before intervention, M = 4.0 and SD=0.60, while after intervention, M = 4.3 and SD = 0.64. The intervention was t(71) = −4.14 and p = 0.000. The results suggest that when exposed to the use of AWE, participants perceived that the suggestions made by AWE are more helpful.

The third paired−samples T−test was conducted on Item 3 which asks whether AWE gives good explanations about errors. Significant difference was indicated as before intervention, M = 3.9 and SD = 0.75 while after intervention, M = 4.2 and SD = 0.74. The intervention was, t(71) = −2.81 and p = 0.006, suggesting that intervention does have an impact on participants; specifically, when exposed to the use

of AWE, participants perceived that AWE gives good explanations on errors made.

The last paired−samples T−test conducted for Theme 1 was on participants' perception on AWE helping participants understand grammar rules (Item 4). There was a significant difference before (M = 3.9, SD = 0.68) and after (M = 4.3, SD = 0.73) with the intervention; t(71) = −5.04, p = 0.000. These results suggest that intervention does have an impact on participants. Specifically, the results suggest that when exposed to the use of AWE, participants perceive AWE helps them understand grammar rules better.

The results from the T−test confirms the H1 for RO2 in which there is a significant difference in respondents' perceptions of AWE assistance in students' writing learning before and after the intervention. Here, AWE is perceived to provide good offers in helping students to improve their writing as it helps provide good feedbacks, suggestions, explanation, and help understand grammar better.

2) Theme 2—Perception on AWE on its usefulness and use in writing

The first paired−samples T−test conducted for Theme 2 was on Item 5 pertaining to the advantage of using AWE in writing class. A significance difference can be seen in before intervention (M = 4.0, SD = 0.6) and after intervention (M = 4.4, SD = 0.66) where the intervention was t(71) = −3.90 and p =0.000. These results suggest that the intervention does have an impact on participants. Specifically, the results suggest that students perceived that they get the advantage of using AWE in writing class.

The paired−samples T−test conducted on Item 6 was pertaining to the ease of use of AWE, especially in writing class. Before intervention, M=4.1 and SD=0.66, and after intervention, M=4.4 while SD=0.66. The intervention was t(71) = −3.70 and p = 0.000, suggesting that the intervention does lead to participants perceiving that AWE is easy to use, especially in writing class.

The next paired−samples T−test for this theme conducted was on Item 7 "If AWE is a better choice for evaluating participants' writing compared to human evaluator?". The result of this test also reveals that there was a significant difference between before intervention (M = 3.5, SD = 0.80) and after intervention (M = 4.0, SD = 0.80). The intervention; t(71) = −4.80 and p = 0.000. These results suggest that after exposure to AWE, participants perceive that AWE is a better choice for evaluating their writing compared to human evaluator.

The paired−samples T−test conducted on Item 8 was aimed to compare participants' preference of AWE over human evaluator before and after the intervention. The difference between before and after intervention was significant where before intervention, M = 3.4, SD = 1.00, and after intervention, M = 3.7, SD = 1.00, with the intervention, t(71) = −3.06, p = 0.003. The results suggest that after exposure to AWE, participants prefer AWE over human evaluator.

The results from the T−test confirms the H1 for RO2 in which there is a significant difference in respondents' perceptions of AWE assistance in students' writing learning before and after the intervention. Here, AWE is perceived to provide advantage to students in writing as it is easy to use and is perceived as a better choice compared to human evaluator hence is more preferrable.

IV. Discussion

People who worked on the internet and wrote in English for daily business recognized the emergence of AWE, especially Grammarly. Being provided online and heavily advertised through social media, AWE has focused on enhancing writing and has expanded its capabilities to provide machine automated aid in language use, along with translation software, online voice dubbing, and other services.

The study's findings analyze English language students' awareness, understanding, and expectations of AWE's benefits in enhancing their writing learning process. Overall, the majority of participants provided positive feedback on their self−perceived proficiency in writing before and after the intervention period. It is also depicted that AWE is preferred to human as writing assistance and evaluator. The study successfully achieved its primary purpose of demonstrating the improvement in participants' writing skills over a 10−week period through the use of Grammarly.

Addressing the first research objective (RO1), the results of this study indicate that AWE does have an impact on participants' perception of their writing ability. Specifically, the results suggest that exposure to the use of AWE leads to a more positive perception of writing proficiency. Nine out of ten analyzed items have shown a significant difference in participants' perception of their writing proficiency before and after the intervention. However, the item related to "practicing making notes or trying to remember feedback" did not show any significant difference before and after the intervention.

The lack of significance in Item 10 may be attributed to the nature of the writing product itself, which is a finished and ready−to−submit piece. Therefore, there is no practical reason for students to take notes or try to remember feedback for the same piece of writing [37].

Addressing the second research objective (RO2), AWE is perceived to offer valuable assistance in helping students improve their writing skills. It provides feedback, suggestions, explanations, and helps improve understanding of grammar. Additionally, AWE is considered advantageous over human evaluators due to its ease of use and availability. AWE is always accessible and provides constant, comprehensive, and fast feedback, which effectively aids in improving students' writing. Moreover, its availability with just one click makes it preferable to a human evaluator.

V. Conclusion

The purpose of this study was to explore the effectiveness of AWE, particularly Grammarly, among university students, as well as their perceptions of its implementation. The findings from the study indicate that Grammarly was successful in improving vocabulary usage, language use, spelling, and punctuation in students' writing. The implementation of Grammarly was viewed favorably by the participants in the survey.

Building upon the results of the study, it is important to note that the use of technology can support the process of ESL writing assessment in this digital learning environment. However, the traditional evaluation methods, including educators explaining and assisting students in organizing their writing and developing meaningful content, would still

be necessary due to several reasons such as poor internet connectivity which posed challenges for students while utilizing the software. On the other hand, Grammarly has the potential to alter the meaning of writing hence, the combination of Grammarly and traditional approaches is crucial for achieving comprehensive ESL writing evaluation results.

Since Grammarly has proven effective in reducing errors related to vocabulary use, language use (grammar), and writing styles (spelling and punctuation), it allows for educators to have more time to focus on helping students revise their writing content and format. However, it is important to also acknowledge the contributions of educators, as they play a vital role in supporting students with varying levels of English proficiency. Educators should consider the appropriate method of integrating Grammarly into writing classes for ESL students.

Additionally, since there are limited studies on the integration of AWE software, further research on other AWE tools is warranted. It is also recommended to conduct similar studies in different educational environments to contribute to the broader field of study. These findings have the potential to benefit students, educators, and writing instructors who are interested in incorporating Grammarly into their ESL writing lessons.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Azman Che Mat and Nurul Ajleaa contributed in the literature section regarding AWE. The authors teamed up in the conceptualization and data collection phases. Luqmanul was responsible for the analysis, pre−writing and the completion of the writing while Azman and Nurul Ajleaa were responsible for the additional information required for the study. All authors had approved the final version.

References

[1] C. Yang, "How Chinese beginning writers learn English writing: A survey of writing strategies," *Journal of Educational and Social Research*, vol. 3, no. 1, pp. 9–18, 2013.

[2] V. L. Fei and J. Phua, "Teaching writing with language feedback technology," *Computers and Composition*, vol. 54, pp. 1–18, 2019.

[3] Maharani, M. M. "Graphic Organizers to Improve Students' Writing on Recount Paragraphs," *Metathesis: Journal of English Language, Literature, and Teaching*, vol. 2, no. 2, 2018, pp. 211–221. https://doi.org/10.31002/metathesis.v2i2.942

[4] J. M. V. Faller, "Grammarly investigation into EFL writing issues involving omani learners," *International Journal of Language & Linguistics*, vol. 5, no. 3, 2018. https://doi.org/10.30845/ijll.v5n316

[5] B. Ghonsooly and S. Shalchy, "Cultural intelligence and writing ability: Delving into fluency, accuracy and complexity," *Novitas−ROYAL (Research on Youth and Language)*, vol. 7, no. 2, 2013.

[6] A. L. Saglam and A. Y. Duman, "Exploring student perceptions of source−based writing assessment in a Turkish EAP context," *Novitas−ROYAL (Research on Youth and Language)*, vol. 14, no. 1, 2020, pp. 25–38.

[7] Ü. Üstünbaş, "In−class or online: Writing self−efficacy and instructional method dilemma," *Novitas−ROYAL (Research on Youth and Language)*, vol. 17, no. 1, 2023, pp. 16–29.

[8] C. Yang, "How Chinese beginning writers learn English writing: A survey of writing strategies," *Journal of Educational and Social Research*, vol. 3, no. 1, 2013, pp. 9–18. doi: 10.5901/jesr.2013.v3n1p9

[9] F. V. Lim and J. Phua, "Teaching writing with language feedback technology," *Computers and Composition*, vol. 54, 2019, 102518.

[10] T. N. Fitria, "Error analysis found in students' writing composition of simple future tense," *ELS Journal on Interdisciplinary Studies in Humanities*, vol. 1, no. 3, 2018, pp. 240–251.

[11] I. Perdana and M. Farida, "Online grammar checkers and their use for EFL writing," *Journal of English Teaching, Applied Linguistics, and Literatures (JETALL)*, vol. 2, no. 2, 2019, pp. 67–76. https://doi.org/10.20527/jetall.v2i2.7332

[12] S. P.−L. Sim, H. P.−K. Sim, and C.−S. Quah, "Online learning: A post COVID−19 alternative pedagogy for university students," *Asian Journal of University Education*, vol. 16 no. 4, pp. 137–151, 2021.

[13] R. Agustin and S. Wulandari, "The analysis of grammatical errors on students' essay writing by using Grammarly," *Jurnal Pendidikan Bahasa Inggris Proficiency*, vol. 4, no. 1, 2022, pp. 39–46. https://doi.org/10.32503/proficiency.v4i1.2247

[14] D. Grimes and M. Warschauer, "Utility in a fallible tool: A multi−site case study of automated writing evaluation," *The Journal of Technology, Learning, and Assessment*, vol. 8, no. 6, 2010, pp. 1–44.

[15] Z. Qiang, "An experimental research on applying automated essay scoring system to college English writing course," *International Journal of English Language Teaching*, vol. 1, no. 2, 2014, pp. 35–41. https://doi.org/10.5430/ijelt.v1n2p35

[16] P. Ware, "Automated writing assessment," in J. I. Liontas & M. Delli Carpini (Eds.), *The TESOL Encyclopedia of English Language Teaching*, pp. 1–7, 2018. https://doi.org/10.1002/9781118784235.eelt0543

[17] O. Zawacki−Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education—where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, 2019.

[18] T. N. Fitria, "Error analysis found in students' writing composition in simple past tense of recount text," *ENGLISH FRANCA: Academic Journal of English Language and Education*, vol. 4, no. 2, 2020, p. 141. https://doi.org/10.29240/ef.v4i2.1154

[19] T. Mammadova, *Teaching Grammar to a Grammar−Free Generation*, Cambridge Scholars Publishing, 2019.

[20] T. N. Fitria, "Grammarly as a teachers' alternative in evaluating Non−EFL students writings," *Leksema: Jurnal Bahasa dan Sastra*, vol. 6, no. 2, 2021, pp. 141–152. doi: 10.22515/ljbs.v6i2.3957

[21] R. O'Neill and A. M. T. Russell, "Stop! Grammar time university students' perceptions of the automated feedback program Grammarly," *Australasian Journal of Educational Technology*, vol. 35, no. 1, 2019, pp. 42–56.

[22] M. A. Ghufron and F. Rosyida, "The role of Grammarly in assessing English as a Foreign Language (EFL) writing," *Lingua Cultura*, vol. 12, no. 4, 2018, pp. 395–4. https://doi.org/10.21512/lc.v12i4.4582

[23] M. S. Soegiyarto, R. A. Putri, and S. D. Saputra, *The Importance of Getting Automated Grammar Feedback via Grammarly, for Increasing Students' English Language Proficiency*.

[24] D. Nuro'azah, "The effectiveness of Grammarly checker toward students' writing quality," *Fakultas Tarbiyah dan Ilmu Keguruan, English*, 2019.

[25] M. Nova, "Utilising Grammarly in evaluating academic writing: A narrative research EFL students' experience," *Journal of English Education and Applied Linguistic*, 2018.

[26] G. L. Parra and S. X. Calero, "Automated writing evaluation tools in improvement of the writing skill," *International Journal of Instruction*, vol. 12, no. 2, 2019, pp. 209–226. https://doi.org/10.29333/iji.2019.12214a

[27] N. A. Zinkevich and T. V. Ledeneva, "Using Grammarly to enhance students' academic writing skills," *Professional Discourse & Communication*, vol. 3, no. 4, 2021, pp. 51–63. https://doi.org/10.24833/2687−0126−2021−3−4−51−63

[28] M. Rahimi and L. J. Zhang, "Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language," *Reading and Writing*, vol. 32, no. 3, 2018, pp. 761–786.

[29] D. I. Hanauer, C. L. Sheridan, and K. Englander, *Linguistic Injustice in the Writing of Research Articles in English as a Second Language: Data from Taiwanese and Mexican Researchers*.

[30] O. Atabek, "Development and validation of digital writing scale for preservice teachers," *Novitas−ROYAL (Research on Youth and Language)*, vol. 14, no. 2, 2020, pp. 119–139.

[31] G. Genç and D. Köksal, "Foreign language learning effort and use of digital media among digital natives: A case study from an urban secondary school," *Novitas−ROYAL (Research on Youth and Language)*, vol. 15, no. 1, 2021, pp. 17–37.

[32] R. Theodora. "Students' perception of self−assessment and the use of self−assessment in English writing skill," *In KOLITA 14: Konferensi Linguistik Tahunan Atma Jaya Keempat Belas*, 2016, pp. 217–221.

[33] J. Gao, "Exploring the feedback quality of an automated writing evaluation system Pigai," *International Journal of Emerging Technologies in Learning*, vol. 11, no. 16, 2021, p. 322. https://doi.org/10.3991/ijet.v16i11.19657

[34] Z. Zhang, "Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions," *Assessing Writing*, vol. 43, 2020, pp. 1–14.

[35] X. Lu *et al*., "An empirical study on the artificial intelligence writing evaluation system in China CET," *Big Data*, vol. 7, no. 2, 2019, pp. 121–129. https://doi.org/10.1089/big.2018.0151

[36] S. Zhang, "Review of automated writing evaluation systems," *Journal of China Computer−Assisted Language Learning*, vol. 1, no. 1, 2021, pp. 170–176. https://doi.org/10.1515/jccall−2021−2007

[37] E. Tavakoli *et al*., "Formative Assessment of Writing (FAoW): A confirmatory factor structure study," *International Journal of Assessment Tools in Education*, vol. 6, no. 3, 2019, pp. 344–−361. DOI: 10.21449/ijate.544277

Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).