



doi <https://doi.org/10.58256/4fd2qt78>

Research Article

Section: Literature, Linguistics & Criticism



Published in Nairobi, Kenya by
Royallite Global

Volume 5, Issue 2, 2024

Article Information

Submitted: 29th January 2024

Accepted: 5th March 2024

Published: 7th April 2024

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

Additional information is available
at the end of the article:

To read the paper online, please scan
this QR code:



How to Cite:

Khasawneh, M. A. S. (2024).
Investigating the impact of automated
instruments used for assessing the
writing skill: Perspectives of language
e-learners. *Research Journal in
Advanced Humanities*, 5(2). <https://doi.org/10.58256/4fd2qt78>

Investigating the impact of automated instruments used for assessing the writing skill: Perspectives of language e-learners

Mohamad Ahmad Saleem Khasawneh

Special Education Department, King Khalid University, Saudi Arabia

Email: mkhasawneh@kku.edu.sa

<https://orcid.org/0000-0002-1390-3765>

Abstract

The main goal of using Automated Writing Evaluation (AWE) instruments is to help language learners learn more quickly and more effectively. In addition to assigning writing-related assignments, these technologies give language learners a suitable forum where they can receive general comments on their written pieces. In the same vein as the benefits of using AWE instruments, this study aims to comprehend foreign language e-learners' perspectives regarding the impact of these tools in their learning environment. Sixty-seven (67) research participants, who are currently e-learning foreign languages at King Khalid University, participated in an online survey to achieve this fundamental goal. Their answers to the questions were compiled and presented using a quantitative method, and they serve as the foundation for the research results. Moreover, the data distribution was precisely computed using a descriptive statistics table. Nonetheless, the study's findings demonstrate that AWE tools come with suitable measurements that non-native English speakers might utilize to improve their writing skills. This is predicated on the scoring propensity and feedback propensity—two crucial aspects of these instruments. Additionally, the study shows that e-learners occasionally run into difficulties when using these resources, which are mostly remediable by the programmers of these applications as well as language teachers. Lastly, to achieve a balance in the appropriate integration and use of AWE tools in the language educational setting, this research recommends using these tools in addition to human feedback and instruction.

Keywords: AWE Instruments, e-learners, foreign language learning, writing skills



© 2024 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY-NC-SA) license.

Public Interest Statement

This study analyzes the differing opinions of language e-learners to remedy this issue. Additionally, the research findings will be based on these viewpoints. In addition to assigning writing-related assignments, these technologies give language learners a suitable forum where they can receive general comments on their written pieces. In the same vein as the benefits of using AWE instruments, this study aims to comprehend foreign language e-learners' perspectives regarding the impact of these tools in their learning environment.

Introduction

For a long time, researchers and educators have used technology to support language instruction and learning. As Ghanizadeh, Razavi, and Jahedizadeh (2015) argue, it has been noticed that using technology in the classroom is just as beneficial as using traditional classroom methods. According to Shadieff, Hwang, and Liu (2018), due to their limited technological knowledge, learners may encounter various difficulties when utilizing technology. Students will probably waste their time and have a bad opinion of technology if they do not receive timely and pertinent support. It is, therefore, necessary to give prompt responses if technical difficulties arise. To use technology for learning more successfully, learners must also be given adequate time to become proficient with it and understand both its benefits and drawbacks.

When implementing technologically assisted learning procedures, learner competency levels must also be taken into account. For instance, learners should be allowed to choose the learning material that best suits their level of skill (Türk & Erçetin, 2014). If not, learning materials may be overly simple or excessively complex, making them unsuitable for learning. The adaptive caption filtering method, which was presented by Hsu (2015), was successful in raising students' listening comprehension because it allowed them to choose from a variety of informational options at varying degrees of difficulty.

The features and application of Automatic Writing Evaluation (AWE) systems have been extensively studied by researchers in the writing profession. These researchers' main goal is to determine how much these tools are used as instructional aids to help language learners become more proficient writers. According to Saricaoglu & Bilki (2021), the creation of Page Essay Grade (PEG) in the 1960s in the US is where AWE first emerged. It appears that the AWE tool is the PEG software's most recent version.

According to studies by Wilson & Roscoe (2020), Aysel & Zeynep (2021), Hyland (2018), and Link et al. (2022), AWE tools help students learn languages and develop their writing abilities. Furthermore, Al-Inbari et al. (2022) found that this technology has a dual purpose: on the one hand, it assists students in developing their writing abilities, and on the other hand, it relieves professors of the laborious process of assessing lengthy student essays, which takes a lot of time.

Aysel & Zeynep (2021) state that the screening scores and constructive feedback function are the two fundamental and all-inclusive features that the AWE tools provide to language learners. These roles were also referred to as the assessment and help functions by Cheng & Cheng (2018). The AWE tools' assessment function involves screening students' written assignments, or essays, and assigning grades following predetermined standards. Although the help feature gives users useful feedback on how well they wrote, a great deal of research has been done to determine how accurate AWE scoring is by comparing the relationships between computer-generated and human raters' scores. AWE technologies like the Intelligent Essay Assessor, IntelliMetric, and e-rater show excellent results based on the measures used, according to Hyland (2018). However in classroom settings, where the content of student writing is assumed to retain a higher value than in standardized exams, the interrater's reliability between humans and computers will be lowered, Keith (2003). On the other hand, the bulk of these studies have mostly concentrated on the

examination of large-scale standardized tests.

One other important issue to be aware of is that AWE software is vulnerable to fraud. Hockly (2019) asserts that skilled writers are capable of fooling automated writing evaluation (AWE) programs and receiving relatively high scores on complex but pointless writing.

2. Review of the Literature

A: Synopsis of Automated Writing Evaluation (AWE) Instruments

Increased awareness of the value of written communication, along with budgetary restrictions and time constraints for precise and trustworthy human assessment and criticism, has made it more important than ever to develop rapid appraisal processes. This has led to the development of Automated Writing Evaluation (AWE) systems. Still, Saricaoglu & Bilki (2021) linked the creation of Page Essay Grade (PEG) in the US during the 1960s as the source of AWE. Over time, the PEG program has been used to grade student writings, detect grammatical problems, and forecast college-level essay scores that are comparable to those of human raters Aysel & Rod (2020). According to Wang (2023), there was a significant alteration to this software throughout the 1990s. Shermis & Hamner (2013) claimed that this breakthrough involved the combination of statistical modeling, natural language processing, and computational linguistics, which resulted in the creation of advanced AWE systems.

The AWE tools are currently made to accommodate a variety of writing assessment goals. According to Yao (2021), one of the AWE systems, the Intelligent Essay Assessor™ (IEA), uses the Latent Semantic Analysis (LSA) method to generate a score. By offering semantics feedback, this method broadens the use of automated scoring, Feng, H.-H. (2015). AWE systems, or automated writing evaluation, support a process in which language learners write repeatedly, get feedback, and edit their work. Some of the well-known tools are PEG Writing (formerly called MI Write), Vantage Learning's IntelliMetric, and Educational Testing Service's e-Rater. Additionally, many other AWE instruments are used in the research industry and may be purchased commercially, according to a study conducted in 2016 by Allen and Perret. Similarly, Strobl et al. (2019) found a large number of about 90 automatic writing evaluators after doing an extensive systematic evaluation.

In general, using AWEs can make it easier to provide multiple cycles of feedback, which gives students the chance to correct mechanical errors and basic organizational or structural flaws before turning in their work to their teacher. Following the dissemination of information, educators might devote their mental energy to providing critical analysis of the content (Link et al., 2020; Wilson and Czik, 2016).

The Application of Computerized Assessment Instruments in Education

Natural language processing powers the scoring algorithms used by Automated Writing Evaluators (AWEs). For open-ended responses, these algorithms produce formative feedback in addition to summative numerical ratings (Chen & Cheng, 2008). Nonetheless, several well-known AWE programs for usage in academic settings have been identified by Wilson and Roscoe (2020), Yao (2021), Aysel & Wilson (2021), and others. These include Criterion, Holt Online Essay Scoring, Writing Roadmap and Write to Learn, and My Access. The two most well-known AWE systems in the K–12 market, according to Joshua and Rod's research (2020), are MY Access ("MA" or "My Access") from Vantage Learning, Inc. and Criterion from a branch of the Educational Testing Service. The academic work states that every piece of software comes with a scoring engine, a unique editing tool that fixes mechanical, grammatical, and spelling errors, in addition to several assistance tools including rubrics, visual organizers, dictionaries, and example essays.

As part of the AWE implementation in the classroom, students are given a specific writing prompt within the software. Students use computers to produce essays in turn. They can type them into a text box or use a word processor to copy and paste the essay together with any reference resources and editing tools

that are offered. Students who submit essays online are graded and provided with feedback right away. Regarding this input, Miranty & Widiati (2021) said that the differences in feedback kinds and quality depend on the specific system under consideration. According to a 2004 study by Attali and Burstein, My Access uses standardized narrative feedback templates that are customized based on the grade level, score, and genre. It's interesting to note that every seventh-grader who receives a 3 on a persuasive essay receives the same suggestions for development. Criterion, on the other hand, provides focused but limited feedback based on the discourse analysis of each essay that is evaluated. It does this by posing questions or making observations about the presence or absence of elements like thesis statements, supporting ideas, and conclusions, among other things. According to Yao (2021), Criterion is better than other AWE applications since it can gather user data and generate detailed results including personality feedback assessments, score analysis reports, and complete outcomes statements. However, these reports can include the total number of words or lexemes as well as the number of errors within particular error categories. Chen & Cheng (2008) go on to say that Criterion generates input about stylistic aspects, use, and sentence structure.

One major advantage of using AWE in the classroom is that it can make students more engaged, which helps them become more proficient in the language. Attali and Burstein (2004) suggest that monitoring students' responses to information obtained from an Automated Writing Evaluation (AWE) feedback system can enhance their level of involvement. For example, after their written works are evaluated and graded, language learners receive critical criticism. After receiving the aforementioned input, it is the individual's responsibility to participate in the modification process. At this point, the student begins the process of editing his or her draft, making the necessary changes, and then turning in the updated paper.

B: Cons of Using AWE Tools in a Language Learning Environment

The use of automatic Writing Evaluation (AWE) technologies, which offer automatic assessment and commentary on written content, has increased in academic settings. AWE tools employ algorithms and preset criteria to examine various aspects of written communication, including the correctness of grammar and punctuation, as well as coherence and structure. The accuracy of the AWE system's representation of human raters has been a source of concern for several academics, including Wilson and Czik (2016), Wang and Wang (2015), and Khoii & Doroudian (2013), despite the system's many advantages. However, Link et al. (2014) have argued that correlational research is typically employed to tackle this problem.

According to research by Shermis et al. (2002), there is a substantial correlation between AWE scores and human scores in a variety of testing settings. The study also reported that in the context of an English placement test, the application of PEG technology produced a significantly higher correlation coefficient.83 about human raters, above the correlation coefficient of .71 noted in human inter-rater dependability. In the context of the TOEFL iBT, Enright and Quinlan (2010) have also calculated agreement indices for ratings given by two human raters and by an e-rater in addition to one human rater. E-rater has proven to be a reliable additional tool to human ratings, displaying higher concurrence indices. These strong correlations support the use and interpretation of scores produced by systems in specific testing contexts.

The vulnerability of AWE software to deception is a significant worry regarding the use of AWE technologies. Hockly (2019) proposed that skilled writers are capable of fooling automated writing evaluation (AWE) systems and receiving relatively high scores on complex but pointless pieces. Chen & Cheng (2008) state that even with poor content, an essay can still receive a good score if it is long and uses certain lexico-grammatical elements that the scoring systems value. As a result, instead of making a sincere attempt to improve their writing abilities, pupils could devise ways to get around these methods. In support of this claim, Yao (2021) pointed out that understanding the specific criteria and computational models used by automated writing evaluation (AWE) software is a critical component that allows proficient authors to

outsmart the software. By altering their compositions to meet the predetermined standards of the program, skilled writers can manipulate their scores. It is conceivable that people could place greater weight on the lexicon, sentence form, and grammar than on deeper factors like thinking, content, and analytical abilities. As such, their compositions may demonstrate a strong grasp of technical abilities but may lack depth and coherence.

Nonetheless, several in-depth research has looked into the significance of the AWE instruments' feedback report. Yao (2021) claimed that the main focus of AWE tools is on grammatical errors and following preset rules, like syntax, orthography, and punctuation; as a result, the more complex facets of writing, like developing substance, logical consistency, or persuasive discourse, are neglected. As a result, students could receive feedback that focuses mostly on small mistakes and ignores important aspects of their writing. Hockly (2019) also talked about the viewpoints of students and their receptivity to AWE criticism. The study indicates that pupils may respond to and prefer different types of feedback. Some students might react more favorably to comments from peers or real teachers than from the AWE technology.

In another study by Wang (2015), the researcher looked at how students perceived using Criterion® to learn English writing as well as the three services it offers (scoring, diagnostic feedback, and writing aid tools). To confirm if the tool was beneficial for student revision, the researcher also computed the revision rates for errors in the students' writing from their initial to final submissions and thoroughly examined the diagnostic feedback messages. The following were the three primary findings. First, the data show that while most students valued the instantaneous scoring and use error analysis, many also expressed dissatisfaction with the scoring rubric, style error analysis, and writing support tool "Plan." Additionally, in a writing class, the majority of students favored the machine score in conjunction with the teacher's explanations. The capacity to turn in 15 drafts in 18 weeks was the program's greatest asset, as reported by the participating students. They also felt that using Criterion® enhanced their English writing skills, albeit they did not credit the Criterion® program directly for the improvement. It's possible that students' advances were caused by their extensive drafting and editing, multiple instances of receiving immediate responses from teachers and machines, and numerous instances of receiving both types of input. One benefit of this study, then, is that it reinforces the conclusions of earlier research (Dyson & Freedman, 1990; Flower & Hayes, 1981), which imply that writing several drafts of a single essay is a necessary but insufficient condition for writing improvement, provided the writer does not receive explicit guidance on how to fix their mistakes and strengthen their writing abilities. The finding that the incorporation of teacher-student mentoring can successfully offset Criterion®'s shortcomings is another contribution made by this research. Students' concerns about the program's shortcomings, such as ambiguous remarks or inaccurate diagnostic signals, were recorded, and all of the participants received one-on-one tutoring and advice on how to write better. Through this kind of advice and counseling, the teacher showed her interest in her students and gave them confidence that she would be there for them anytime they needed help understanding the machine's recommendations. Second, the study validated students' beliefs that certain diagnostic feedback signals from Criterion® were helpful for student revision while others were not, based on the revision rates of errors in students' written work. These results also align with those of other research (Burstein & Marcu, 2003; Chen, Chiu, & Liao, 2009; Grimes, 2008; Ware, 2011), which demonstrated that AWE algorithms were better at identifying surface problems in students' compositions, such as spelling or grammar mistakes.

Furthermore, in a study by Zhu, Liu, and Lee (2020), the researchers examined a feedback system integrated with a science curriculum module that students learn in the classroom, making use of automated scoring and natural language processing techniques. To enhance their written scientific argumentation skills, the researchers looked at how students responded to the automated feedback and how performance improvements are connected to the adjustments made possible by the feedback system. Additionally,

the generic and contextualized forms of feedback were contrasted in this study. It was discovered that contextualized feedback was more effective than generic feedback at promoting performance improvements because it included more content-specific suggestions for modifications. In comparison to students who received generic criticism, those who received contextualized feedback saw comparable score increases in fewer rounds of modifications.

3.1 Problem Statement

It appears that several studies have looked into the applicability and consequences of using AWE technologies in the classroom. On the other hand, very few studies have looked at how language learners—particularly online learners—perceive the impact of using these resources on the improvement of their writing skills. This study intends to analyze the differing opinions of language e-learners to remedy this issue. Additionally, the research findings will be based on these viewpoints.

3.2 Research Questions

1. What impact do AWE tools have on learning foreign languages?
2. How much do the grading and feedback elements of AWE tools encourage and involve online learners in their language learning environment?
3. What difficulties do online learners have when utilizing these resources in the classroom?

4. Research Methods

A Study Methodology

The current study used a quantitative research design to examine how foreign language e-learners' use of AWE tools influences their learning. In addition, the study collected research participants' comments on the study's issue via the use of an online questionnaire.

A. Targeted Sample

Sixty-eight (68) e-learners from King Khalid University who are presently enrolled in foreign language courses make up the research participants. On the other hand, these individuals were chosen at random using online discussion boards. A questionnaire with pertinent research questions for the current study was also given to these participants. However, it is important to remember that there are differences in the amount of demographic data, such as gender and age, when it comes to the fundamental details about the participants. The study recognized the significance of taking ethics into account when gathering data. To protect the subjects' privacy, informed consent was obtained before any data was collected. Nevertheless, Table 1 presents the participants' complete distribution according to relevant demographic information.

Table 1. Demographic Data

Category	Variable	Frequency	Percentage
Gender	Female	50	50.0%
	Male	50	50%
Age	18-25 years	38	55.88%
	26-35 years	17	25%
	36-45 years	9	13.23 %
	46 years above	4	5.88 %

Educational Background	Bachelors' degree	47	69.11 %
	Masters' degree	17	25 %
	Ph.D.	4	5.88 %
Experience with AWE Tools	Yes	42	61.76 %
	No	26	38.23 %

Relevant demographic data on the research participants is displayed in the above table. The following is a summary of this table:

- i. Of the sixty-eight (68) research participants, half are female and half are men.
- ii. The bulk of participants, or roughly 55%, are between the ages of 18 and 35. However, only a smaller portion of the population—5.88 %—is over the age of 46.
- iii. Additionally, among the research participants, 47 (69.11 %) hold a bachelor’s degree, 17 (25%) have a master’s degree, and 4 (5.88%) have a PhD.
- iv. On the other hand, there are 26 (38.23%) research participants without any experience with AWE technologies, compared to 42 (61.76 %) who have.

A. Research instruments

An online questionnaire was utilized in this investigation. As opposed to this, the research participants’ replies on this five-point Likert-scale questionnaire span from strongly agree to strongly disagree. In addition, there are four parts to the survey instrument’s content. Important demographic data regarding the research participants is included in the first section. Following this, the second section presents the research participants’ perspectives regarding the impact of AWE tools on foreign language e-learners’ learning process. The third relates to the second research topic, which looks into how much the AWE tools engage and encourage e-learners. Nevertheless, the last section aims to present the participants’ perspectives on the difficulties posed by using these instruments.

B. Data Analysis Strategy

After creating a questionnaire using a Likert scale to gather relic data for the study, the data was presented and examined using a descriptive statistical method. To determine the distribution of the research participants’ responses, this method’s application involves using a descriptive statistics table with key central tendencies like mean and standard deviation.

5. Analysis and Presentation of Data

This section focuses on investigating the three fundamental research questions. However, the following table of descriptive statistics is consistent with the initial study query.

A. How do AWE instruments affect learning a foreign language?

The first study question’s item is as follows:

- 1. Language learners’ writing skills increase when they consistently use AWE tools.
- 2. Is it possible for language learners who utilize AWE tools to outperform their peers in writing assignments?
- 3. When using AWE tools for writing-based tasks, there is a noticeable decrease in errors.
- 4. AWE instruments support overall fluency in foreign languages.

Table 2. AWE Tools’ Effect on Learning Foreign Languages.

Question Variable	SA	A	N	D	SD	Mean	S.D
Q1	32.53%	45.34%	16.19%	3.44%	2.5%	3.22	1.03
Q2	25.2%	41.24%	23.23%	7.42%	2.91%	3.17	0.97
Q3	28.02%	48.24%	17.4%	4.21%	2.13%	3.32	0.97
Q4	33.23%	42.14%	18.42%	4.31%	1.9%	3.41	0.95

The statistical opinions of research participants regarding the potential improvement of foreign language learners’ writing skills through consistent use of AWE tools are presented in the above table. On the other hand, the following summarizes how the above table should be understood;

- i. With a mean score of 3.22 and a standard deviation of 1.03, the research participants highly appreciate the value of the AWE tools in enhancing their writing ability. However, the lower percentages of 16.19% neutral 3.44% disagree and 2.5% strongly disagree in the responses show that there are positive responses among the participants.
 - ii. However, there is a moderate degree of agreement among participants regarding the question of whether language learners who utilize AWE tools in their writing-intensive work outperform their peers. But the mean score, which is 3.17, makes this clear. Furthermore, over 50% of the participants confirmed this, with the remaining, relatively small, amount of participants—less than 15%—rejecting this claim.
 - iii. Regarding the second question item, a sizable portion of participants (76.26%) confirmed that using AWE tools significantly reduces mistakes in writing-related tasks. Less than 20% of individuals were indifferent to the statement, while the remaining percentage disagreed.
 - iv. Participants also agree on the final survey question in a moderate amount. More than half of the participants confirmed that using the AWE tools enhances not only writing but also other language skills. The remaining participants disagreed with this statement, with 18.42% of them staying neutral.
- B. What is the degree to which the scoring and feedback functions of AWE tools encourage and involve online learners in their language learning setting?

Listed below are the survey questions from the second research question;

- i. Language learners’ motivation is increased by AWE score and feedback elements.
- ii. The AWE scoring and feedback tools support language learners’ active participation in the language learning process.
- iii. Language learners are happier with the way the AWE tools grade and provide feedback.

Table 3. The utilization of AMAZING resources to inspire and involve language learners

Question Variable	SA	A	N	D	SD	Mean	S.D
Q1	26.17%	44.86%	17.76%	9.35%	1.87%	4.17	0.85
Q2	32.71%	37.38%	18.69%	7.48%	3.74%	3.73	0.84
Q3	25.5%	24.04%	27.36%	14.08%	9.02%	3.12	0.80

Regarding the second study question, the distribution of research participants' responses is shown in the above table. Nevertheless, the table is explained below;

- i. Significantly, over 50% of participants confirmed that the AWE grading and feedback outcomes improve motivation and help language learners participate actively in the language learning process. A smaller portion of participants—less than 20%—rejected this claim, nevertheless.
- ii. However, there are differing opinions on whether language learners are happy with the AWE tools' grading and feedback features. On the other hand, 49.54% of participants agreed with this argument, 27.36% were undecided, and the rest participants did not.

C. How difficult is it for online learners to use these resources in the classroom?

The following outlines the substance of the third research survey question:

- i. The AWE tools' feedback and grading criteria take into account the demands and learning preferences of the students.
- ii. AWE technologies handle the social and collaborative aspects of writing assignments.
- iii. While acknowledging different writing styles, AWE technologies fail to appropriately capture cultural nuances.
- iv. Language learners can manipulate the AWE system for higher grades.
- v. Following set standards is given more weight by AWE than encouraging creativity and originality in writing.

Table 4. Constraints on utilizing AWE tools in a classroom

Question Variables	SA	A	N	D	SD	Mean	S.D
Q1	12.12%	15.45%	25.36%	32.08%	15.99%	2.72	1.02
Q2	15.89%	25.23%	27.10%	19.63%	12.15%	3.09	0.867
Q3	24.30%	19.63%	17.76%	30.84%	7.47%	2.78	0.98
Q4	19.63%	32.71%	21.50%	15.88%	10.28%	2.59	1.15
Q5	19.64%	23.36%	25.23%	18.69%	13.08%	2.91	0.98

The research participants' answers to the final research question are shown in the above table. Meanwhile, the following interpretation of the table is made:

- i. The first item's mean score of 2.72 indicates the degree of disagreement in the responses. The majority of participants (i.e., less than 50%) appear to disagree with the claim that the AWE tools' grading and feedback outcomes correspond to students' requirements and learning styles. A sizable portion of respondents—25.36 percent—were neutral, but the remaining percentage disagreed with this assertion.
- ii. The research participants hold varying opinions on how AWE tools address the social and collaborative aspects of writing. A mere 40% of participants confirmed this statement, whereas 26.17% expressed no opinion. Less than 32% of respondents also strongly disagreed with this statement, indicating that they don't agree with the collaborative and social components of AWE tools in the writing processes.
- iii. Participants feel that AWE tools don't appropriately reflect writing styles and cultural differences in the third item. Despite this, 43.93% of respondents agreed with this, while 17.76% had no opinion. Additionally, 38.31% of participants question the ability of AWE tools to capture writing styles and

cultural nuances.

- iv. In the meantime, more than half of the participants confirmed that language learners can rig AWE systems to get higher grades in the fourth item. While others expressed no opinion, 36.16% disagreed with this assertion.
- v. Lastly, the final item demonstrates that participants believe AWE tools foster conformity to rules rather than promoting uniqueness and imaginative writing. Less than 33% of respondents are apathetic, although over 40% of respondents agreed with this. Although AWE tools highlight creativity and originality in writing, 25.23% of participants disagreed.

6.1 Data Analysis Discussion

The research participants who are e-learners of foreign languages have responded, as shown in the data analysis above. On the usability and efficacy of the Automated Writing Evaluation (AWE) tools, however, the participants shared their opinions. The respondents' answers served as the foundation for the data employed in this study, which was then utilized to address the three main research questions.

The first study question examines the effectiveness of the AWE, and the respondents offered their differing perspectives on it. However, the information collected from this study question showed that language learners' writing skills can be enhanced by routinely using AWE tools. Language learners are likely to consciously or unconsciously build competent writing skills given the many stages that entail the usage of these instruments. Writing with AWE tools, feedback and revision, additional revision, and teacher help or intervention are some of the pertinent processes that Hockily (2019) identified as being connected to the usage of AWE tools in the classroom. However, Yao (2021) went on to say that the reason instructors should be involved in helping students with their writing is that they can carefully examine the feedback reports generated by the Automated Writing Evaluation (AWE) tools to understand the strengths and weaknesses of each student's written task. In addition, providing more detail on the explanations provided by the AWE instruments would help students understand the suggestions and apply them to their written assignments.

Most of the respondents also expressed without hesitation that using AWE tools improves performance on any writing-related work. Additionally, the result supports the claims made by Aysel & Rod (2020) and Wang (2023). Theoretical works proposed that feedback outcomes help students identify the areas of their writing proficiency that require further development. As a result, students who take note of these comments are more likely to scan their work for mistakes that could potentially be related to them, which enables them to generate high-quality writing. Concurrently, the third thing corresponds with the second item as well. AWE technologies help prevent common errors in writing-related tasks, according to the majority of research participants. However, Wang (2023) highlighted how AWE tools have an algorithm built in that makes it simple to suggest typical linguistic errors. Because of this, when learners utilize these tools, they are made aware of this mistake and make an effort to avoid it when completing writing-related tasks. It's also evident that AWE tools do more than just help language learners become better writers. The procedures or stages involved in using these tools often lead to the development of other language abilities in learners, including speaking, listening, and reading.

Additionally, most of the answers to the second research question show that language learners benefit from the AWE tools because they are motivated and encouraged to write, which helps them become more proficient writers. Participants' opinions of how satisfied they are with the feedback and grading results of the AWE tools, however, differ somewhat. According to Yao (2021), the tool's incapacity to grade pupils above and beyond its preset criteria is likely the root cause of the various attitudes of learners regarding the AWE tool's score and feedback outcomes. Furthermore, this adequacy also includes the inability to recognize and take into account the writing styles of different people.

The final research aims to investigate the difficulties that students have when using AWE technologies. The difficulty of these tools to meet the demands and learning styles of students, their lack of social and collaborative writing skill elements, their inability to capture cultural nuances, and the system's tendency toward manipulation are just a few of their drawbacks. Guo (2020) pointed out that the AWE tools are already set up to a specific standard, which may not match the demands or learning styles of language learners in particular. This has to do with the student's learning styles and needs. To prevent this problem, Dang et al. (2013) recommended the adoption of alternative methods for feedback and evaluation. However, in this case, language teachers offer students not only more feedback but also individualized counseling based on their needs. The makers of these AWE tools have suggested adapting pertinent tactics, which is another essential answer offered by Wilson & Roscoe (2020). These tactics include real-time collaboration, adaptive feedback and guidance, multimodal features, and user-content design and feedback.

The absence of social and collaborative components in writing practice is another drawback that is consistent with Chen & Cheng's (2008) results. Peer engagement is one of the elements that promotes quicker language acquisition, according to Dang et al. (2013). Therefore, when they solely engage with the AWE system, language learners often become bored more quickly. However, to guarantee the thorough development of writing proficiency, Delsa et al. (2021) endorsed the notion of integrating peer review sessions in addition to the AWE tools.

7. Conclusion

The impact of Automated Writing Evaluation (AWE) technologies was investigated in this study. Nevertheless, a careful examination of these resources reveals that the AWE method helps language learners not only with writing but also with other language proficiency areas including speaking, listening, and reading. Furthermore, language learners can be motivated and actively involved in their learning process because of the effectiveness of the AWE tools. This is nonetheless predicated on the two fundamental characteristics that AWE tools have: feedback results and scoring. The scoring tool grades students' written work in real time. On the other hand, language learners can receive feedback via the feedback feature regarding any errors or mistakes that require correction. There is a common concern about the difficulties in using these technologies, in addition to their benefits. The inability to accurately record and take into account the various writing styles of individuals, the tools' propensity for manipulation, etc., are some of these issues. Nonetheless, it is up to the tool developers and teachers to find a solution to these problems. The AWE system can be effectively enhanced and utilized to support language learners in enhancing their writing competence by implementing pertinent tactics. In the meanwhile, this study suggests integrating these tools with human evaluation and guidance to establish a balance in the appropriate integration and application of AWE technologies in the context of language teaching.

Funding: This research was funded by the Deanship of Scientific Research at King Khalid University for funding this work through Large Research Groups under grant number (RGP.2 / 108 /45).

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Research Groups under grant number (RGP.2 / 108 /45).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Aysel Saricaoglu, Zeynep & Bilkii (2021). *Voluntary use of automated writing evaluation by content course students*. Cambridge University Press, 33, 265-265
- Burstein, J., Chodorow, M. & Leacock, C. (2003) Criterion SM online essay evaluation: An application for automated evaluation of student essays. In Riedl, J. & Hill, R. (eds.), *Proceedings of the Fifteenth Innovative Applications of Artificial Intelligence Conference*. Menlo Park: AAAI Press, 3–10.
- Burstein, J., & Marcu, D. (2003). A machine learning approach for identification of thesis and conclusion statements in student essays. *Computers and the Humanities*, 37(4), 455–467.
- Chapelle, C. A. & Voss, E. (2016) 20 years of technology and language assessment in language learning & technology. *Language Learning & Technology*, 20(2), 116–128.
- Chen, C.-F. E. & Cheng, W.-Y. E. (2008) Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112.
- Cotos, E., Huffman, S., & Link, S. (2020). Understanding graduate writers' interaction with and impact of the Research Writing Tutor during revision. *Journal of Writing Research*, 12(1), 187-232.
- Council of Europe (2001) Common European framework of reference for languages: Learning, teaching, assessment. Strasbourg: Language Policy Unit.
- Dang, T. K. A., Nguyen, H. T. M. & Le, T. T. T. (2013) The impacts of globalization on EFL teacher education through English as a medium of instruction: An example from Vietnam. *Current Issues in Language Planning*, 14(1), 52–72.
- Delsa Miranty, Utami Widiati. Pegem (2021). Automated Writing Evaluation (AWE) in Higher Education: Indonesian EFL Students' Perceptions about Grammarly Use across Student Cohorts. *Journal of Education and Instruction*, 11(4), 2021 (pp. 126-137).
- Don Yao (2021), Automated Writing Evaluation for ESL Learners: A Case Study of Pigai System. *The Journal of Asia TEFL*, 18(3), Fall 2021, 949-958.
- Dyson, A. H., & Freedman, S. W. (1990). *On teaching writing: A review of the literature*. Berkeley, CA: National Center for the Study of Writing.
- El-Ebyary, K. & Windeatt, S. (2010) The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121–142.
- Feng, H.-H. (2015) Designing, implementing, and evaluating an automated writing evaluation tool for improving EFL graduate students' abstract writing: A case in Taiwan. Iowa State University, unpublished Ph.D. <https://lib.dr.iastate.edu/etd/14824/>
- Ferris, D. R., Liu, H., Sinha, A. & Senna, M. (2013) Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22(3), 307–329.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.
- Ghanizadeh, A., Razavi, A., & Jahedizadeh, S. (2015). Technology-enhanced language learning (TELL): A review of resources and upshots. *International Letters of Chemistry, Physics and Astronomy*, 54, 73-87.
- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Learning and Assessment*, 8(6).
- Grimes, D. C. (2008). Middle school use of automated writing evaluation: A multi-site case study (Unpublished doctoral dissertation). University of California, Irvine, CA.
- Guo, L. (2020). A study of the application of Pigai software in teaching university-level English writing: Taking a freshman class in Chongqing as an example. *Theory and Practice in Language Studies*,

- Hibert, A. I. (2019). Systematic literature review of automated writing evaluation as a formative learning tool. In *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14* (pp. 199-212). Springer International Publishing.
- Hockly, N. (2019). Automated writing evaluation.
- Ho, O. K., Yao, D., & Kunnan, A. J. (2021). An analysis of Macau's joint admission examination-English. *The Journal of Asia TEFL*, 18(1), 208-222.
- Hsu, C. K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL*, 27(1), 84-103
- Joshua Wilson, Cristina Ahrendt, Emily A. Fudge, Alexandria Gaysha Beard, Chares MacArthur (2021). Elementary teachers' perceptions of automated feedback and automated scoring: Transforming the teaching and learning of writing using automated writing evaluation. *Computer and Education*, Volume 168.
- Koh, C. (2016) Translating motivational theory into the application of information technology in the classroom. In Liu, W. C., Wang, J. C. K. & Ryan, R. M. (eds.), *Building autonomous learners: Perspectives from research and practice using self-determination theory*. New York: Springer, 245–258.
- Koltovskaia, S. (2020) Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44: 1–12.
- Lai, C., Yeung, Y. & Hu, J. (2016) University student and teacher perceptions of teacher roles in promoting autonomous language learning with technology outside the classroom. *Computer Assisted Language Learning*, 29(4): 703–723.
- Lai, Y.-h. (2010) Which do students prefer to evaluate their essays: Peers or computer programs? *British Journal of Educational Technology*, 41(3): 432–454.
- Laviolette, E., Polio, C. & Kahng, J. (2015) The accuracy of computer-assisted feedback and students' responses to it. *Language Learning & Technology*, 19(2): 50–68.
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49, 1-7
- Leki, I. (2006) "You cannot ignore": L2 graduate students' response to discipline-based written feedback. In Hyland, K. & Hyland, F. (eds.), *Feedback in second language writing: Contexts and issues*. New York: Cambridge University Press, 266–285.
- Link, S., Dursun, A., Karakaya, K., & Hegelheimer, V. (2014). Towards better ESL practices for implementing automated writing evaluation. *Calico Journal*, 31(3), n3.
- Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 35(4), 605-634.
- Li, J., Link, S. & Hegelheimer, V. (2015) Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27: 1-18
- Li, R., Meng, Z., Tian, M., Zhang, Z., Ni, C. & Xiao, W. (2019) Examining EFL learners' antecedents on the adoption of automated writing evaluation in China. *Computer Assisted Language Learning*, 32(7):
- Li, Z., Feng, H.-H. & Saricaoglu, A. (2017) The short-term and long-term effects of AWE feedback on ESL students' development of grammatical accuracy. *CALICO Journal*, 34(3): 355–375.
- Liao, H.-C. (2016) Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, 70(3): 308–319.
- Mackey, A. & Polio, C. (2009) Introduction. In Mackey, A. & Polio, C. (eds.), *Multiple perspectives on*

- interaction: Second language research in honor of Susan M. Gass. New York: Routledge, 1–10.
- Norris, J. & Ortega, L. (2003) Defining and measuring SLA. In Doughty, C. J. & Long, M. H. (eds.), *Handbook of second language acquisition*. Malden: Blackwell, 717–761.
- Parra G, L., & Calero S, X. (2019). Automated writing evaluation tools in the improvement of writing skills. *International Journal of Instruction*, 12(2), 209-226.
- Palermo, C., & Wilson, J. (2020). Implementing automated writing evaluation in different instructional contexts: A mixed-methods study. *Journal of Writing Research*, 12(1), 63-108.
- Plonsky, L. & Ziegler, N. (2016) The CALL–SLA interface: Insights from a second-order synthesis. *Language Learning & Technology*, 20(2), 17–37.
- Polio, C. (2012) The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21(4): 375–389.
- Quinlan, T., Higgins, D. & Wolff, S. (2009) Evaluating the construct-coverage of the e-rater scoring engine. Princeton: Educational Testing Service.
- Saricaoglu, A. (2019) The impact of automated feedback on L2 learners' written causal explanations. *ReCALL*, 31(2), 189–203.
- Saricaoglu, A., & Bilki, Z. (2021). Voluntary use of automated writing evaluation by content course students. *ReCALL*, 33(3), 265-277.
- Shadiev, R., Hwang, W. Y., & Liu, T. Y. (2018). A study of the use of wearable devices for healthy and enjoyable English as a foreign language learning in authentic contexts. *Journal of Educational Technology & Society*, 21(4), 217-231.
- Türk, E., & Erçetin, G. (2014). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning*, 27(1), 1-25.
- Wang, P.-l. (2013) Can automated writing evaluation programs help students improve their English writing? *International Journal of Applied Linguistics & English Literature*, 2(1), 6–1.
- Ware, P. (2011). Computer-generated feedback on student writing. *TESOL Quarterly*, 45(4), 769–774.
- Wilson, J. & Czik, A. (2016) Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers & Education*, 100, 94–109
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58, 87-125
- Zhang, S. (2021). Review of automated writing evaluation systems. *Journal of China Computer-Assisted Language Learning*, 1(1), 170-176.
- Zhu, M., Liu, O. L., & Lee, H. S. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education*, 143, 103668.