Artificial intelligence in education: Effects of using integrative automated writing evaluation programs on honing academic writing instruction

Bantalem Derseh Wale
College of Social Sciences and Humanities, Injibara University, Ethiopia
*Corresponding Author: bantalemd@gmail.com

ABSTRACT

Automated writing evaluation (AWE), which is the result of educational artificial intelligence technology, is a process of scoring and evaluating learners' written texts automatically. The current study examined the effects of using integrative AWE programs on honing academic writing instruction. It also assessed students' perceptions towards using these programs. A quasi-experimental pretest-posttest two-group design was used. Test, questionnaire, focus group discussion, and teacher diary were used to collect data from 92 randomly selected participants. The experimental group students learned writing skills with Writerly and Google Docs in integration, but the control group students learned through the conventional paper and pencil feedback system. When the quantitative data were analyzed through independent samples T-test and descriptive statistics, the qualitative data were analyzed thematically. The findings revealed that using the integrated AWE programs honed academic writing instruction because there was a statistical difference between the experimental and control groups in their academic writing performance. Hence, students who learned using the integrated AWE programs honed their academic writing performance because they were able to produce essays that addressed task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. However, students who learned through the conventional method were less effective in producing quality essays. Besides, the findings also discovered that the experimental group students had positive perceptions towards using the AWE as mentioned above programs because they found the programs interesting, effective, goal-oriented, and supportive. Consequently, this study recommends researchers, curriculum designers, instructional material designers, teachers, and students pay due attention to integrated AWE programs.

Keywords: artificial intelligence, automated writing evaluation, Google Docs, writerly, academic writing perception

INTRODUCTION

The field of education has always been greatly influenced by the rapid development of educational technology. Automated Writing Evaluation (AWE) can be defined as a process of scoring and evaluating learners' written work automatically through computer programs (Shermis & Burstein, 2003). It is an e-program that traced its origin to the 1960s in the United States with the evolution of Page Essay Grade (PEG) which works based on a collection of previously rated writing samples (Page, 2003 and Zhang, 2021). As a result of the advancement of educational Artificial Intelligence (AI) technology in general, and the enhancement of natural language processing and intelligent language tutoring systems in particular, the design of AWE programs has been improving rapidly since the mid1990s (Lane et al., 2013; Cotos, 2014; Chen & Cheng, 2018, and Jingxin & Razali, 2020).

AWE is designed to provide instant computer-generated scores for a submitted essay along with diagnostic feedback (Chen & Cheng, 2018, and Zhang, 2021). Hence, most AWE programs offer complementary writing instruction and give diagnostic feedback in terms of organization, mechanics, grammar, diction, and language use (Ariyanto et al, 2021). Several AWE programs including IntelliMetric, E-rater, The Intelligent Essay Assessor, Pigai, iTEST, iWrite, Project Essay Grade, My Access!, Criterion, Holt Online Essay Scoring, Writing Roadmap and Write to Learn, Grammarly, Wordtune, Writerly, and Google Docs have been launched to evaluate and enhance written texts (Zhang, 2021 and Zhanga & Huang, 2020). These AWE programs often provide instant scores along with corrective feedback in various aspects of writing and can be used for both formative and summative assessment purposes. Hence, they serve as a writing assistant tool beyond assessment for they have editing features that work on grammar, diction, spelling, and style checkers. Accordingly, students use AWE programs to write and revise their essays in a self-regulated learning environment (Chen & Cheng, 2018, and Saricaoglu & Bilki, 2021).

Previous studies including Ranalli, et al. (2016), Zhanga & Huang (2020), Jingxin & Razali (2020), and Asratie et al. (2022) found that the use of AWE systems positively affects students' writing performance. For example, Ranalli, et al. (2016) reported that using AWE decreased learners' writing errors. Zhanga & Huang (2020) and Jingxin & Razali (2020) showed the use of AWE programs improved college EFL learners' writing skills as well as increased their learner autonomy. It provides written corrective feedback in aspects of vocabulary, sentence structure and organization, and content, based on a large corpus of standard English.

Among several AWE programs, newly developed AI writing technologies like "Writerly" and "Google Docs" are becoming popular in writing instruction because they have user-friendly features like content outline, sentence booster, and collaborative writing (Seyyedrezaie, 2016). Writerly is an online writing application that utilizes artificial intelligence to generate texts based on students' input. Hence, it combines automation, integration, and artificial intelligence to meet students' academic needs quickly and efficiently. It takes students' ideas and thoughts and makes them more fluid by adjusting the organization, diction, content, tone, and style of the text (Saricaoglu & Bilki, 2021). This online software program has academic and nonacademic platforms including editing, advertising, marketing, sales, e-commerce, social media, website, academic, recruiting, real estate, and long form. Of all these platforms, editing, academic, and long-form are the major writing features of the program that enable students to enhance their writing quality. Specifically, the 'Editing' service of this software program enables students to use the 'Grammar Improver', 'Elaborator', 'Synonym', 'Sentence Booster', and 'Convincing Bullet Points. Besides, the 'Academic' feature of the program incorporates 'Elaborator', 'Essay Assistant', and 'Essay Outline' which are helpful to students enhance the quality of their written texts. Using the 'Essay Outline' platform, students type their essay writing topic so that the software generates the contents of the essay that they can incorporate into their essay. In addition, in the 'Essay Assistant' and 'Elaborator' sub-platforms, the students drop their paragraphs in the given space so that the software enhances the contents and ways of expression. Furthermore, this software program translates students' essays from and into English language which helps students to better express their ideas. Therefore, via these specific features, the writerly software program provides several global feedbacks such as content outline, idea development, and sentence booster to hone academic writing instruction. On the other hand, Google Docs which is an online collaborative writing tool allows students to collaboratively produce written texts, receive peer and teacher feedback, and revise and edit essays synchronously which is suitable to be applied in academic writing instruction (Seyyedrezaie 2016).

Conjointly, the integrative use of Writerly and Google Docs in writing classrooms enhanced academic writing instruction because students use both software programs jointly in producing essays. Particularly, the writerly program lies in its ability to provide continuous and constructive feedback to the writers as well as opportunities to review and revise their work until they are satisfied with it. In doing so, learners revise their drafts by referring to the formative automated feedback. In addition, it also helps students to overcome the dilemma caused by nervousness and shyness, which usually happens in EFL students from African countries when

they interact with their teacher and peers in conventional face-to-face classrooms (Jingxin & Razali, 2020, and Seyyedrezaie, 2016). Besides, via the Google Docs software program online peer feedback and teacher feedback have been performed, and as a result, students' writing performance has been enhanced.

However, while several AWE programs are researched, Writerly and Google Docs are less studied, particularly in the current research context, Ethiopia. Hence, although the use of Google Docs in foreign language writing instruction has been previously studied (Seyyedrezaie, 2016), there are no adequate studies conducted on the integrated use of Writerly and Google Docs in academic writing instruction. These AWE programs were chosen for the purpose of this study because the researcher had prior experience with them. Therefore, the aim of the present study was to investigate the effects of using integrative automated writing evaluation programs on honing academic writing instruction in Ethiopia. In addition, this study was designed to assess students' perceptions towards using integrative automated writing evaluation programs in academic writing instruction.

Writing is a difficult skill to acquire in language learning for students to articulate their ideas with correct written language (Fan & Ma, 2022, and Birhan, 2018). It is a challenging task for EFL learners to outline main points, organize ideas, revise unity and coherence, and edit for mechanics (Wale & Bogale, 2021). In order to help students enhance their writing performance, the provision of corrective feedback on students' writing becomes necessary, but it is challenging for writing teachers. Besides, Zhang, 2021; Derseh, 2020, and McNamara, Crossley, Roscoe, Allen, & Dai, 2015) also mentioned that conventional writing instruction takes an inordinate amount of teacher time to score students' essays and provide subsequent feedback to the students. Foltz et al. (1999), and Wale & Bishaw (2020) also stated that while writing is an essential part of the educational process, many teachers find it difficult to incorporate a large number of writing assignments in their courses due to the effort required to provide corrective feedback. Chen (2022) and Lee (2019) stressed that marking students' writing is a highly challenging job, and teachers usually have to devote a substantial amount of time to giving corrective feedback to their students. Reading and correcting students' writing is time-consuming for teachers. Especially in developing countries where there are more than 50 students in a class, asking students to write more means teachers have to devote extended periods of time to assessing and giving comments on students' written works (Wang, 2013).

A promising solution to this challenge is using integrative AWE programs that employ artificial intelligence to evaluate essays and offer instant corrective feedback. With the rapid development of educational technology, studies on AWE programs have been gaining more ground in EFL writing instruction, particularly due to its potential to give continuous, corrective feedback on students' written texts (Chen, 2022, and Jingxin & Razali, 2020). Zhang (2021) also stated that using AWE has the advantages of time and cost saving, efficiency, and a learner-centered feedback process. Automated writing feedback can clearly reduce a teacher's workload by providing detailed feedback, and students can receive feedback immediately after submitting their writing. The feedback process becomes learner-centered because learners can conduct self-assessments online. AWE systems provide opportunities for students to write online, receive timely feedback, and revise their writing accordingly in an iterative way. In such a context, learner agency plays an important role, as learners comprehend feedback information, make judgments for further improvement, and take responsibility for their learning.

In examining the effectiveness of the AWE program on improving learners' writing ability, (Jingxin & Razali, 2020; Wang, 2013, and LinHuang, 2010) stated that the use of the Criterion and CorrectEnglish AWE programs enhanced students' writing performance in the aspect of linguistic accuracy and grammar in Taiwan. Hence, using the aforementioned AWE programs, the students corrected their grammatical errors including fragments, subject-verb disagreement, run-on sentences, and ill-formed verbs. Ibid stressed that the feature of immediate feedback of AWE makes learning more efficient and interesting. On the contrary, critics of AWE argued that the validity of AWE programs is doubtful. For example, Chen & Cheng, 2018, and Attali & Burstein, 2006) distrust the ability of computers to "read" texts and evaluate the quality of writing because computers are unable to understand meaning in the way humans do. They also doubt the

value of writing to a machine rather than to a real audience, since no genuine, meaningful communication is likely to be carried out between the writer and the machine. Moreover, they worry whether AWE led students to focus only on surface features and formulaic patterns without giving sufficient attention to meaning in writing their essays.

All in all, although most previous research findings on educational artificial intelligence AWE systems were progressive and fruitful, they are few, and their findings seem to be contradictory, inconclusive, and insufficient in terms of causational empirical studies. Besides, most previous studies on AWE have been conducted by a survey or an interview on psychometric evaluations of its validity; however, studies on the effectiveness of AWE programs in writing instruction as a pedagogical tool are limited. Therefore, there was a conspicuous research gap to support that the integrative AWE programs, Writerly and Google Docs, enhance EFL students' writing performance. Accordingly, the current study employed an experimental research design to make the previous research findings more comprehensive. Besides, though many previous studies have tended to examine AWE's accuracy and validity in scoring essays, little research has paid attention to the effectiveness of giving and receiving automated feedback on improving EFL learners' writing performance. Hence, little research has been conducted on the effectiveness of AWE feedback on improving learners' writing performance, and students' perceptions towards using integrative AWE programs.

On the other hand, most of the aforementioned studies were conducted out of the EFL context that exclude the setting where the current study was conducted. In other terms, though Writerly and Google Docs programs are being used by writing teachers and students worldwide, the effectiveness of using these integrative AWE programs in enhancing academic writing instruction was not adequately investigated in the Ethiopian context. Besides, studies examining Ethiopian students' perceptions towards using these integrative AWE programs were also scant. Hence, as far as the researcher's reading is concerned, there are no adequate studies that examined the effects of using integrative AWE programs on honing EFL students' academic writing instruction in the Ethiopian context.

In light of these concerns, the purpose of this study was to examine the potential of artificial intelligence in English language education in general and in providing written feedback to hone EFL academic writing instruction. Particularly, it focused on the effects of using integrative automated writing evaluation programs (particularly, Writerly and Google Docs) on honing academic writing instruction in the Ethiopian context. Accordingly, the current study addressed the following two research questions: 1) What are the effects of using integrative automated writing evaluation programs on honing academic writing instruction? 2) What are students' perceptions towards using integrative automated writing evaluation programs?

METHOD

The aim of this study was to examine the potential of artificial intelligence in English language education in general and in honing students' academic writing performance in particular. Hence, it focused on the effects of using integrative automated writing evaluation programs (Writerly and Google Docs) on honing academic writing instruction. Besides, it also aimed to assess students' perceptions towards using integrative automated writing evaluation programs. Accordingly, the study employed a quasi-experimental research design which entails a pretest and post-test with two groups of participants escorted by a mixed research approach (Creswell, 1994).

The participants were 92 third-year English as a Foreign Language Department students who were taking 'Advanced Writing Skills' course with the researcher at Injibara University, Ethiopia. They were in two groups (sections 'A and C'), and the sections were randomly selected from a total of four sections, 'A' to 'D' containing 194 third-year in-service students who were following their education in the Ethiopian Summer season. The participants had formally studied English language subject for 14 years: 12 years in school and two years in the University. Their English language proficiency was approximately at the upper-intermediate level for they had taken "Communicative English Language Skills", "Basic Writing Skills" and "Intermediate Writing Skills" courses in their freshman and sophomore years.

The data were collected through essay writing test, questionnaire, focus group discussion, and teacher diary. Accordingly, two sample IELTS essay writing proficiency tests (pretest and posttest) were used to assess the participants' writing performance. The pretest was administered before the intervention to both experimental and control group students to evaluate their existing academic writing performance. The test reads: *"The internet has transformed lives and economies but it is turning the world into a global village. Soon everybody will think and behave in the same way. To what extent do you agree or disagree with this opinion? Present your argument in an essay using over 250 words arguing for or against this idea"*. Besides, the post-test was administered after the intervention to both experimental and control groups of students to determine whether the intervention made a difference in the students' academic writing performance. It reads: *"The growing number of smokers is putting a strain on the health care system in an effort to deal with the health issues involved. Some people think that the best way to deal with this problem is to legally prohibit smoking cigarettes. To what extent do you agree or disagree with this opinion? Present your argument in an essay using over 250 words arguing for or against this idea"*. Both the pretest and post-test students' essays were scored by two experienced writing teachers using the British Council IELTS writing task-2 descriptors that incorporate task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy (British Council, 2018). The Pearson's correlation's Inter-rater reliability of the two raters was 0.8, reliable.

The second data gathering instrument was questionnaire needed to gather data on students' perception towards using integrative AWE programs (Writerly and Google Docs). It was designed using Google Forms incorporating 18 five-point Likert-scale type questions and four open-ended type questions, and administered via Google Forms to the experimental group students after the intervention. The reliability of the questionnaire was 0.79 as calculated by Cronbach's α, which indicated that it was reliable.

Thirdly, focus group discussion was also used to collect data on students' perception towards using integrative AWE programs, and the development of students' academic writing performance. The discussion was made with 12 experimental group students using eight thought-provoking questions that revolved around the effectiveness of AWE programs, Writerly, and Google Docs.

Finally, teacher-diary which is the teacher's daily classroom note was also used to collect data on the effectiveness of integrative AWE programs in honing academic writing instruction and the enhancement of students' writing performance.

In the data collection process, first, the data gathering instruments were designed and piloted. Then, a sample IELTS writing test was administered to all four sections of English language and literature department students to select a homogeneous group of participants. Based on the diagnostic test, two sections of students ('A' and 'C') were found homogeneous and selected to be the participants of the study. Section 'A' students were assigned as experimental group, and section 'C' students were assigned as control group randomly. Following the group assignment, the IELTS sample essay writing pre-test was administered to both experimental and control groups to get baseline data on the students' academic writing performance: task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. After collecting the baseline data through the pretest, the intervention was conducted for eight weeks.

During the intervention, the experimental group students (Section 'A') learned argumentative essay writing using the integrative AWE programs, Writerly and Google Docs, while the control group students (Section 'C') learned argumentative writing with the conventional method. Hence, the experimental group students followed their essay writing instruction in the English Language Improvement Centre (ELIC) where the Writerly and Google Docs writing tools were installed on its desktop computers. Besides, these online software programs were also installed on the experimental group students' smartphones. Before the intervention was started, the experimental group students received short-term training on how Writerly and Google Docs work. Hence, the teacher assisted students in how to use the Writerly-specific features like Editing, Essay Outline, Sentence Booster, Grammar Improver, Essay Assistant, Elaborator, Synonym, and Convincing Bullet Points. Accordingly, using this online

software program the students outlined their essay contents, boosted sentences, worked on grammar and synonyms, elaborated their expressions, and revised and edited their essays several times. Besides, the students also wrote argumentative essays collaboratively and gained peer and teacher feedback using Google Docs to improve their written texts and hone their academic writing performance. Hence, the students were categorized into groups to share feedback on each other's argumentative essays via Google Docs. The group members were required to read and provide corrective feedback on the other groups members' argumentative essays based on IELTS writing task-2 descriptors that incorporated task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. Following the peer feedback, the students received the teacher's feedback on the correctness of the peer feedback that might be incorrect or incomplete so that they could receive balanced and better feedback on their draft essays. Moreover, the teacher also provided detailed corrective feedback on the quality of the essays based on the IELTS above writing task-2 descriptors. When the students had difficulties in revising and editing their draft essays based on the automated feedback, the teacher assisted them on how to revise and edit their draft essays via Writerly and Google Docs. On the other hand, the control group students also learned argumentative essay writing face-to-face through the conventional paper and pencil feedback system and with their printed teaching material that incorporated the theoretical aspects of argumentative essay writing.

When the intervention was conducted, the teacher diary was held in both the control and experimental groups. Hence, the teacher has recorded the daily teaching-learning experiences such as the strengths, weaknesses, and impacts of Writerly and Google Docs from the experimental group and the conventional teaching-learning practices of the control group. At the end of the intervention, the sample IELTS essay writing posttest was administered to the experimental and control group students to evaluate the students' academic writing performance. Following the posttest, the questionnaire and the focus group discussion were conducted with the experimental group students to collect data on the effectiveness of using the integrated AWE programs, Writerly and Google Docs, and the students' perception towards using the integrated programs. Finally, the students' essays were assessed by two raters who were experienced writing teachers, while the questionnaire and focus group discussion data were analyzed by the researcher.

The data were analyzed using quantitative and qualitative methods. The quantitative data collected through tests were analyzed through independent samples T-test in SPSS version-25 software program to examine whether there were differences between the experimental and control groups of students' academic writing performance that incorporates task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. Besides, the close-ended questionnaire data collected on students' perceptions were also analyzed quantitatively using descriptive statistics including mean and standard deviation. On the other hand, the qualitative data gathered through open-ended questionnaire, focus group discussion, and teacher-log were thematically analyzed using qualitative data analysis methods and presented coherently based on common themes.

FINDINGS AND DISCUSSION
Findings
Honing academic writing instruction through integrated AWE programs
The results showed that the students honed their academic writing performance through the integrative AWE programs, Writerly and Google Docs. As the descriptive statistics result depicted in Table 1 indicated, both the experimental and control group students had comparable academic writing performance in the pre-test. Hence, while the experimental group students had a mean score of 45.04; the control group students had a 44.78 mean score.

Thus, though it appeared the students' results had some difference, the variance they had is statistically insignificant. Therefore, it can be understood that both the experimental and control group students had comparable academic writing performance prior to taking the treatment. Nonetheless, the experimental group (M=54.68) and control group (M=45.85) students' post-test results mean scores were statistically significant as can be seen in Table 1. Accordingly, based on the variation observed between the pretest and posttest results mean scores, it can be concluded

that the experimental group students who learned the advanced writing skills course with integrated AWE programs showed better enhancement in their academic writing performance over the control group students that learned the course with the conventional paper and pencil feedback system.

Table 1. Descriptive statistics of experimental and control groups

Test	Participants' Group	N	Mean	Std. Deviation	Std. Error Mean
Pre-test	Experimental Group	45	450.444	416.127	.62033
	Control Group	47	447.872	364.117	.53112
Post-test	Experimental Group	45	546.889	477.599	.71196
	Control Group	47	458.511	407.526	.59444

Most essentially, the independent samples t-test was run to understand the differences between the experimental and control group students' pre-test and post-test results. As depicted in Table 2, the Levene's test for equality of variances showed no violations, P=.529 in the pre-test and P=.616 in the post-test. Besides, the pre-test result also indicates that there was no statistically significant difference between the experimental and control groups before the intervention t (90) =0.4000, P>.05, d=0.814). However, the posttest results disclosed that there was a statistically significant difference between the experimental and control groups t (90) =0.253, P<.05, d=0.924).

Table 2. Independent samples t-test of experimental and control groups

	Test	F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Pre-test	Equal variances assumed	0.400	0.529	0.316	90	0.753	0.25721	0.81425
	Equal variances not assumed			0.315	87.287	0.754	0.25721	0.81663
Post-test	Equal variances assumed	0.253	0.616	9.562	90	0.000	883.783	0.92429
	Equal variances not assumed			9.529	86.514	0.000	883.783	0.92750

It reveals that the students who had learned academic writing skills through the integrated AWE programs, Writerly and Google Docs, outperformed in their academic writing performance compared to the students who learned the skills in the conventional face-to-face and paper and pencil method of teaching.

On the other hand, the results gained through the teacher diary also ensured that the students' academic writing performance was honed when they practiced writing using Writerly text-enhancing features and with Google Docs. In the first week of the intervention, the teacher in his diary noted.

> *It is towards the beginning of the training. The students who are using Writerly and Google Docs are facing several challenges in editing, revising, and organizing their essay content through the software programs. However, the other groups who are learning without the AWE programs are writing better essays for they are receiving written feedback face to face with no machine distraction.*

Nevertheless, towards the end of the intervention, the following was recorded in the teachers' diary.

> *Now, the students who are using the AWE programs are producing better essays in terms of task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. They are almost properly using Writerly and Google Docs to outline their essay contents, draft coherent sentences and paragraphs, write essays collaboratively, share peer feedback and teacher feedback, and correct the*

comments accordingly. Their essays are by far better compared to their own previous essays. However, the essays that are written by the students who are not using the Writerly and Google Docs programs are not that much improving in their writing compared to their previous performance because their dictions are not good, the language used is not that sound, the paragraphs are not well organized. Even, the students get into confrontation when they write essays in face-to-face groups. They got board to repeatedly revise their draft essays using pen and paper.

In addition, the teacher in his diary recorded: "Using the Writerly and Google Docs reduced my workload as a teacher since they provided detailed feedback to the students' essays. They are also time and cost-saving since they generate learner-centered feedback that could be corrected by the students".

Moreover, the experimental group students in their open-ended questionnaire also reported that they had developed their academic writing performance while using Writerly and Google Docs through integration. For instance, one of the students recorded:

I had known something about Google Docs before this time, but the Writerly system was totally new to me. I was challenged to write my first argumentative essay using Writerly. However, after some time, I had written good essays with the help of the software. The Google Docs was also interesting to me because it helped us to write essays together with my group members. I feel that the software programs have supported me to improve my writing skills.

Similarly, another student also wrote:

Writerly is amazing to me. I used to struggle to set my essay contents, boost sentences, organize ideas in paragraphs, revise and edit essays before this training. In this training, the teacher has shown us this technique which is surprising to most of us. All can be done with it! The program gave me several options to boost my sentences and paragraphs. In addition, Google Docs helps us to write paragraphs and essays with friends on a certain title. We do not need to print our essays and submit them to the teacher because he himself can access the essays via Google Docs and give us constructive comments".

Finally, the students in their focus group discussion also ascertained that using Writerly and Google Docs in amalgam enhanced their academic writing performance. Hence, most of the focus group discussion members agreed that they have developed their writing performance due to using the automated software programs in their writing course.

Consequently, it can be generalized that the use of the integrative AWE programs, Writerly and Google Docs, in the advanced writing skills course was effective in honing the students' academic writing performance including achievement, coherence and cohesion, lexical resource, grammatical range and accuracy.

Students' perception toward using integrative AWE programs in writing instruction

The students' perceptions towards using the integrated AWE programs, Writerly and Google Docs were assessed through questionnaire, focus group discussion and teacher-diary. The results showed that the students had positive perceptions towards using the integrative AWE programs for these software programs honed their academic writing performance in terms of achievement, coherence and cohesion, lexical resource, grammatical range and accuracy.

Specifically, the students' questionnaire results indicated that the Writerly and Google Docs programs were interesting, effective, and goal-oriented, and as a result, they enhanced students' writing performance. Table 3 disclosed that the students had (M=4.16; SD=0.737) regarding whether they enjoy writing through Writerly and Google Docs. This result implied that most of the participants enjoyed writing through Writerly and Google Docs. Similarly, most respondents (M=4.20; SD=0.815) would like to use Writerly and Google Docs to write essays. Besides, most of the students (M=4.07; SD=0.780) were interested to use Writerly and Google Docs in their writing instruction. Thus, it implied that the students were interested in using the integrative AWE programs, Writerly and Google Docs.

Table 3. Students' interest in using the integrated AWE programs

No	Items	Mean	Std. Deviation	Std. Error of Mean
1	I enjoy writing through Writerly and Google Docs.	4.16	0.737	0.110
2	I would like to use Writerly and Google Docs to write essays.	4.20	0.815	0.121
3	I was interested in using Writerly and Google Docs.	4.07	0.780	0.116

The results on the students' views on the effectiveness of integrated AWE programs indicated that the use of Writerly and Google Docs was effective. For instance, as depicted in Table 4, the students had (M=4.22; SD=0.823) on whether Writerly took their essay writing idea and made it more fluid by adjusting the organization, diction, content, tone, and style of the text. It can be, therefore, deduced that the students have a positive perception towards Writerly because this software program took their writing idea and made it more fluid by adjusting the organization, diction, content, tone, and style of the text. Likewise, the students viewed (M=4.13; SD=0.757) that Writerly provided them several global feedback such as content outline, idea development, and sentence booster to hone their writing instruction. Hence, this result showed that they positively perceived Writerly since it offered them several global feedbacks. In the same manner, the participants understood (M=4.04; SD=0.767) that Google Docs allowed them to edit their written texts synchronously. In the same way, the students perceived (M=4.20; SD=0.757) that Google Docs enabled them to write essays collaboratively with their colleagues. Lastly, they also thought (M=4.18; SD=0.777) that Google Docs helped them comment their colleagues written works online. Therefore, the results indicated that the students had positive perception towards using Google Docs in their writing instruction because it allowed them to edit written texts synchronously; enabled them to write essays collaboratively with their colleagues, and helped them comment on their colleagues' written works online.

Table 4. Students' view on the effectiveness of integrated AWE programs

No	Items	Mean	Std. Deviation	Std. Error of Mean
1	Writerly took my essay writing idea and made it more fluid by adjusting the organization, diction, content, tone, and style of the text.	4.22	0.823	0.123
2	Writerly provided me with several global feedbacks such as content outline, idea development, and sentence booster to hone writing instruction.	4.13	0.757	0.113
3	Google Docs allowed me to edit my written texts synchronously.	4.04	0.767	0.114
4	Google Docs enabled me to write essays collaboratively with my colleagues.	4.20	0.757	0.113
5	Google Docs helped me comment on my colleagues' written works online.	4.18	0.777	0.116

Table 5 shows that the students viewed the integrative AWE programs were goal-oriented to produce quality written texts. Specifically, the participants had (M=4.22; SD=0.765) about whether Writerly provided them corrective feedback in terms of organization, content, grammar, diction, mechanics, style, and language use. In the same manner, the students perceived (M=4.27; SD=0.863) that Writerly helped them revise and edit their essays in a self-regulated learning environment. They also understood (M=4.16; SD=0.673) that Writerly generated their essay contents that could be incorporated into their essays. Similarly, the participants viewed (M=4.09; SD=0.793) that the Writerly program boosts ways of written expressions. Besides, the students pointed out (M=4.16; SD=0.796) that Writerly translated written texts from and into English helped them to better express ideas. These results uncovered that the students had positive perceptions towards using the AWE software programs, Writerly and Google Docs for the software and were goal-oriented in producing quality written texts. Hence, the Writerly program provides them with corrective feedback; enables them to revise and edit their essays; generates

essay contents that could be incorporated into their essays; boosts ways of written expressions, and translates written texts from and into English that helps them to better express ideas.

Table 5. Students' viewpoint on whether the integrated AWE programs are goal-oriented to better write

No	Items	Mean	Std. Deviation	Std. Error of Mean
1	Writerly provides corrective feedback in terms of organization, content, grammar, diction, mechanics, style, and language use	4.22	0.765	0.114
2	Writerly helped me revise and edit my essays in a self-regulated learning environment.	4.27	0.863	0.129
3	Writerly generates essay content that could be incorporated into essays.	4.16	0.673	0.100
4	The Writerly software boosts ways of written expressions.	4.09	0.793	0.118
5	Writerly translates written texts from and into English which helps to better express ideas.	4.16	0.796	0.119

Table 6 presented, the relevance of integrating Writerly and Google Docs in writing instruction. Accordingly, the students perceived (M=4.11; SD=0.859) that the integrative use of Writerly and Google Docs helped them to overcome the nervousness and shyness that most of them had faced to interact face-to-face. Similarly, the participants viewed (M=4.29; SD=0.626) that using Writerly and Google Docs through integration increased their writing achievement. Likewise, the students understood (M=4.27; SD=0.688) that the use of Writerly and Google Docs in integration improved the coherence and cohesion of their essays. Besides, they thought (M=4.36; SD=0.773) that the integration of Writerly and Google Docs enhanced the lexical resources of their essays. In the same way, the students viewed (M=4.38; SD=0.716) that integrating Writerly and Google Docs together boosted the grammatical range and accuracy of their written texts. The results, therefore, implied that the students' perception towards using the integrative AWE programs, Writerly and Google Docs was positive for the programs helped them overcome their nervousness and shyness, developed the lexical resources of their essays, increased their writing achievement, improved the coherence and cohesion of their written texts, and, and boosted the grammatical range and accuracy of their written texts.

Table 6. The relevance of integrating Writerly and Google Docs in writing instruction

No	Items	Mean	Std. Deviation	Std. Error of Mean
1	The integrative use of Writerly and Google Docs helped me to overcome the nervousness and shyness that I had faced to interact face to face.	4.11	0.859	0.128
2	Using Writerly and Google Docs through integration increased my writing achievement.	4.29	0.626	0.093
3	The use of Writerly and Google Docs in integration improved coherence and cohesion	427	0.688	0.102
4	The integration of Writerly and Google Docs enhanced lexical resources.	4.36	0.773	0.115
5	Integrating Writerly and Google Docs together boosted grammatical range and accuracy.	4.38	0.716	0.107

The students' focus group discussion results also assured that students had positive perceptions towards using integrative AWE software programs, Writerly and Google Docs, for the programs were interesting to use, effective for developing written texts, goal-oriented to produce quality essays, and relevant in writing instructions. Hence, most of the focus group discussion participants agreed that the Writerly program was enjoyable, and they would like to use it for their future academic and career journeys. For instance, one of the focus group discussion participants explained that Writerly has helped her to outline the contents of her essay,

develop major ideas, and boost sentences. Besides, the other participant also reported that Google Docs supported her in editing draft essays, producing paragraphs and essays collaboratively, obtaining comments from friends, and receiving feedback from her writing teacher.

One of the participants in the focus group discussion also voiced:

> *I enjoyed using Writerly to write the essays [because] it was really helpful to translate the essay into Amharic [Ethiopian national language]. I also translated our Amharic written paragraphs into the English language based on the software program. I was surprised when the program gave us interesting essay content after we wrote our essay title on the software. It was funny and friendly.*

In the same manner, the other member of the focus group discussion also reported:

> *The Writerly feedback is helpful and amazing. For example, from what I remember, when I wrote a sentence on Sentence Booster in the Writerly, it totally modified the sentence's content. The new sentence was grammatically correct and written with better word choices. The computer corrected the organization of the essay, the content, grammar errors, and word choice.*

In addition, most of the participants in the focus group discussion agreed that using Writerly and Google Docs in integration was fascinating to them because it enhanced their writing performance in terms of writing task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy which are the qualities of a good essay based on IELTS writing task-2 descriptors. Thus, the results uncovered that the use of Writerly and Google Docs through integration was positively perceived by the students.

Besides, the results gained through the teacher diary also confirmed that most of the students were comfortable with Writerly and Google Docs in producing essays with them. The note recorded in the diary in the middle of the intervention reads:

> *Now it seems that the students have understood how the software programs work. They are entertaining with Writerly when they receive immediate feedback after submitting their unfinished essays to this software program. Today, the students felt happy and appreciated the Writerly program when they themselves ran the feedback process and received software-generated comments on their specific essays.*

Besides, the questionnaire results also confirmed that these online programs were interesting to them to receive and correct draft essays. It unveiled that the students had positively perceived Writerly and Google Docs. The results, therefore, revealed that the students' perception towards the integrated AWE programs, Writerly and Google Docs, was positive, for these AWE programs were interesting, effective, and goal-oriented.

All in all, the results divulged that the integrative use of Writerly and Google Docs programs, which are the products of artificial intelligence, was effective in honing students' academic writing performance that incorporated task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. Besides, the results also uncovered that the students' perceptions towards using these educational artificial intelligence software programs were positive for the programs provided them several correctives feedback which are helpful to producing quality written texts.

Discussion

This study examined the potential of educational artificial intelligence in English language instruction in general and in honing academic writing instruction in particular. Predominantly, it focused on the effects of using the integrative AWE programs, Writerly and Google Docs, on honing academic writing instruction in the Ethiopian context. In addition, it also assessed students' perception towards using integrative AWE programs. The results uncovered that the integrative use of Writerly and Google Docs in advanced writing instruction was effective to honing students' academic writing performance that incorporated task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy. Hence the current study unveiled that the students who had learned writing skills through the aforementioned integrated AWE programs outperformed in their academic writing performance compared to the students who learned the skills face-to-face through the conventional paper and pencil feedback system with

their printed learning material that incorporated the theoretical aspects of writing. This finding is in accordance with Ranalli, et. al, 2016; Seyyedrezaie, 2016, and Zhanga & Huang, 2020 that found AWE provides students with accurate information to target relevant areas of revision, improvement, and learning that help students improve their writing performance.

The present study found that the Writerly AI software program provided students with corrective feedback; enabled them to revise and edit their essays; generated essay contents that could be incorporated into their essays; boosted ways of written expressions, and translated written texts from English into Amharic and visa vice that helped them express their ideas better. This finding bears resemblance to Saricaoglu & Bilki, 2021; Liao, 2016, and Li, Feng & Saricaoglu, 2017 uncovered that AWE feedback had positive effects on the reduction of errors and enriched the quality of student written texts. Besides, the current study showed that while the Writerly program took the students' writing ideas and made them more fluid by adjusting the organization, diction, content, tone, and style of the text, the Google Docs enabled students to edit their written texts synchronously; write essays collaboratively with their colleagues, and comment their colleagues written works online. In line with this result, Aken (2017) discovered that AWE programs enabled learners to collaborate with each other online and receive feedback from both the teacher and colleagues to enhance organization and language usage.

On the contrary, this research finding does not semblance to Warschauer & Grimes, 2008; Ariyanto et al, 2021, and Chen & Cheng, 2018 that discovered since the feedback was predetermined and unable to provide context-sensitive responses involving rich negotiation of meaning, AWE was not useful for content development. For example, Chen & Cheng (2018) discovered that most advanced language learners found AWE unhelpful in producing their written texts because the computer system could not understand contextual meaning and was unable to address their writing problems including coherence and idea development. Additionally, Ariyanto et al. (2021) also reported that teachers overlooked the implementation of AWE in their writing classes for there was a risk that their students only engaged sketchily with the programs by hitting the correction directly.

Besides, the current study also assessed students' perceptions towards using the integrative AWE programs, Writerly and Google Docs. The results discovered that the students had positive perceptions towards using these educational artificial intelligence software programs because they found these online software programs interesting to use, effective to be used in writing instructions, goal-oriented to develop essays, and supportive to overcome their nervousness and shyness faced in face to face conversation.

This finding is in congruence with Wang (2013) and Seyyedrezaie (2016) that showed the pedagogical writing practices with AWE software positively affected students' perceptions when the students used the program to facilitate their drafting and revising process. Specifically, Wang (2013) underscored that AWE programs build students' confidence and motivated them to write more because the programs are stress-free and entertaining to correct feedback online. Similarly, Seyyedrezaie (2016) also indicated that the students had positive attitude towards Google Docs which enabled them to write collaboratively through an entertaining online learning environment. Quite the reverse, this specific finding is in contradiction with Ariyanto et al, 2021; Chen & Cheng, 2018; Cheng, 2006, and Yang, 2004 that revealed students were dissatisfied with the use of AWE in their writing classrooms because the AWE feedback was vague and repetitive. Particularly, Chen & Cheng (2018) specified that students whose language proficiency level were advanced, did not need to circumscribe themselves with automated criteria because they believed that the machine-generated form-focused responses were inconsequential. Overall, the current research findings resemble most previous research findings though they also gainsaid with some other previous discoveries.

CONCLUSION

The current study explored the potential of educational artificial intelligence in English language instruction in providing written corrective feedback to hone academic writing instruction. It, specifically, focused on examining the effects of using integrative AWE programs

namely Writerly and Google Docs on honing academic writing instruction in the Ethiopian context. Besides, it also assessed students' perception towards using integrative AWE programs. The findings of the study revealed that using the integrative AWE programs which are the products of artificial intelligence honed academic writing instruction. In other terms, the integrative use of Writerly and Google Docs in writing instruction was effective in honing students' academic writing performance because these online software programs enabled the experimental group students to produce essays that fulfilled writing task achievement, coherence and cohesion, lexical resource, grammatical range and accuracy that are the essential features of academic writing. On the contrary, the control group students were unable to develop quality essays that satisfied the aforementioned descriptors because the control group students learned writing skills through the conventional paper and pencil feedback system with their printed learning material that incorporated the theoretical aspects of writing.

The findings showed that while the Writerly program took students' writing ideas and made them more fluid, generated essay content outlines, boosted ways of written expressions, and translated written texts from English into Amharic and visa vice that helped students express their ideas better, the Google Docs program also enabled students to write essays collaboratively with their colleagues, and comment their colleagues written works online, receive teacher and peer feedback, and revise and edit their written texts synchronously. The results also disclosed that the students had positive perceptions towards using the integrative AWE programs, Writerly and Google Docs to hone their academic writing performance because these educational artificial intelligence programs were interesting to the students, effective in developing written texts, goal-oriented to produce quality essays, and relevant in writing instructions.

Thus, using the integrated AWE programs in writing instruction is recommended to hone EFL academic writing instruction. Hence, it implies that it is significant to instigate using the integrated AWE programs to hone academic writing instruction and enable students to enhance their writing performance. Besides, writing teachers need to use Writerly and Google Docs in their writing instruction to hone their students' academic writing performance. Correspondingly, writing course material developers should reconsider these integrated AWE programs while developing writing course instructional materials. Moreover, students have to use these integrated AWE programs to enhance their academic writing performance since writing is required in their academic journey and future careers.

However, due to limited time and resources, this study involved a small number of students and focused on only two AWE programs. Nonetheless, it does not mean that the findings of the study were not inclusive if the selected AWE programs were used by the experimental group students during the intervention period. Equally, it does not mean that the intervention period was inadequate because the students practiced writing skills with the integrated AWE programs for two consecutive months. So, it means that the findings of the current research would have been more inclusive if more amount of time and resources had been used for the intervention, a greater number of students participated in the study, and a greater number of AWE programs were integrated into the intervention. Therefore, it is recommended that future studies be conducted with longer duration, more resources, larger sample sizes, and different writing achievements using more advanced AWE programs.

ACKNOWLEDGEMENT

REFERENCES

Aken, A. (2017). An evaluation of assessment-oriented computer-based text analysis paradigms. *Higher Education Research*, 2(4), 111–116. https://doi.-org/10.11648/j.her.20170204.12

Ariyanto, M.S., Mukminatien, N., & Tresnadewi, S. (2021). College students' perceptions of an automated writing evaluation as a supplementary feedback tool in a writing class. *Jurnal Ilmu Pendidikan (JIP), 27*(1), 41-51.

Asratie, M.G., Wale, B.D. & Aylet, Y.T. (2022). Effects of using educational technology tools to enhance EFL students' speaking performance. *Education and Information Technologies.* https://doi.org/10.1007/s10639-022-11562-y

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater v.2. *Journal of Technology, Learning, and Assessment, 4*(3),1-30.

Birhan, A. T. (2018). Effects of mastery learning instruction on engineering students' writing skills development and motivation. *Journal of Language and Education, 4*(4), 20–30. https://doi.org/10.17323/2411-7390-2018-4-4-20-30.

British Council. (2018). *IELTS Task 2 Writing Band Descriptors: Public Version*. British Counsil.

Chen, C.E. & Cheng, W.E. (2018). Beyond the design of automated writing evaluation: pedagogical practices and perceived learning effectiveness in efl writing classes. *Language Learning & Technology, 12*(2), 94-112.

Chen, Z., Chen, W., Jia, J., & Le, H. (2022). Exploring AWE-supported writing process: An activity theory perspective. *Language Learning & Technology, 26*(2), 129–148. https://doi.org/10125/73482

Cheng, W. Y. (2006). *The use of a web-based writing program in college english writing classes in taiwan: a case study of myaccess.* Unpublished Master's thesis. National Kaohsiung First University of Science and Technology, Taiwan.

Cotos, E. (2014). Genre-based automated writing evaluation for l2 research writing: from design to evaluation and enhancement. *Palgrave Macmillan*.

Creswell, J.H. (1994). *Research design: qualitative, quantitative, and mixed methods approaches.* (4th Ed.). SAGE Publications, Inc.

Derseh, B. (2020). Enhancing EFL students' writing performance through inquiry-based learning. *Italian Journal of Educational Research,* XIII:*40*, 138-156. DOI 10.7346/SIRD-012020-P138.

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. https://files.eric.ed.gov/fulltext/EJ936915.pdf https://doi.org/10.6018/ijes/2010/2/119231

Fan, N. & Ma, Y. 2022). The effects of automated writing evaluation (awe) feedback on students' english writing quality: a systematic literature review. *Language Teaching Research Quarterly,* (28), 53-73.

Foltz, P.W., Laham, D., & Landauer, T.K. (1999). The intelligent essay assessor: applications to educational technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning,* 1, 2, http://imej.wfu.edu/articles/1999/2/04/index.asp.

Jingxin, G. & Razali, A.B. (2020). Tapping the potential of pigai automated writing evaluation (awe) program to give feedback on efl writing. *Universal Journal of Educational Research*, *8*(12B): 8334-8343. DOI: 10.13189/ujer.2020.082638

Lane, H.C., Yacef, K., Mostow, J. & Pavlik, P. (2013). *Artificial intelligence in education:* 16th International Conference, Memphis, TN, USA, July 9-13 Proceedings.

Lee, I. (2019). Teacher written corrective feedback: Less is more. *Language Teaching, 52*(4), 524-536. https://doi.org/10.1017/S0261444819000247

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. https://journals.equinoxpub.com/index.php/CALICO/article/view/26382

Liao, H.-C. (2016) Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal, 70*(3): 308–319. https://doi.org/10.1093/elt/ccv058

LinHuang, S. H. (2010). *The exploitation of e-writing in an EFL classroom: Potential and challenges.* Unpublished Master's thesis. I-Shou University, Taiwan.

Page, E. (2003). Project essay grade: PEG. In M. D. Shermis & J. Burstein (Eds.), Automated essay scoring: A cross-disciplinary perspective, 43-54. *Mahwah, NJ: Lawrence Erlbaum Associates.*

Piaget, J. (1973). To understand is to invent: The future of education (G. Roberts, Trans.). NY: Grossman Publishers.

Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2016). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology, 37*(1), 8-28.

Saricaoglu, A., and Bilki, Z. (2021). Voluntary use of automated writing evaluation by content course students. *ReCALL, 33*: 3, 265–277. doi:10.1017/S0958344021000021

Seyyedrezaie, Z.S., Ghonsooly, B., Shahriari, H., & Fatemi, A.H. (2016). A mixed methods analysis of the effect of google docs environment on efl learners' writing performance and causal attributions for success and failure. *Turkish Online Journal of Distance Education-TOJDE, 17*(3), 7.

Shermis, M.D. & Burstein, J. (2003). *Automated essay scoring: A cross-disciplinary perspective,*147-167. Mahwah, NJ: Lawrence Erlbaum.

Wale, B.D. & Bishaw, K.S. (2020). Effects of using inquiry-based learning on EFL students' critical thinking skills. *Asian-Pacific Journal of Second and Foreign Language Education,* 5:9. https://doi.org/10.1186/s40862-020-00090-2

Wale, B.D. & Bogale, Y.N. (2021). Using inquiry-based writing instruction to develop students' academic writing skills. *Asian-Pacific Journal of Second and Foreign Language Education*, 6(4). https://doi.org/10.1186/s40862-020-00108-9

Wang, P. (2013). Can automated writing evaluation programs help students improve their english writing? *International Journal of Applied Linguistics & English Literature, 2*(1).

Yang, N. D. (2004). Using myAccess in EFL writing. *The proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics* (pp. 550-564). Taipei, Taiwan: Ming Chuan University.

Zhang, S. (2021). Review of automated writing evaluation systems. *J. China Comput. Assist. Lang. Learn, 1*(1): 170–176. https://doi.org/10.1515/jccall-2021-2007

Zhanga, L. & Huang, Z. (2020). Effects of an automated writing evaluation system on students' efl writing performance. So, H. J. et al. (Eds.). Proceedings of the 28th International Conference on Computers in Education. *Asia-Pacific Society for Computers in Education.*