Advanced Journal of Interdisciplinary Studies 9(2026) 2568-2593

Journal homepage:www.ajids.dmu.edu.et

Volume 9(2), March, 2026

Effects of Automated Written Corrective Feedback on First-year Undergraduate EFL Students' Paragraph Writing Performance: Wollo University in Focus

Manaye Mesfin[1], Dawit Amogne[2]

[1]Department of English Language and Literature, College of Social Sciences and Humanities, Wollo University, Dessie, Ethiopia [2]Department of English Language and Literature, School of Language Studies, Addis Ababa University, Addis Ababa, Ethiopia

Abstract

The main purpose of this study was to examine the effects of automated written corrective feedback on first-year undergraduate EFL learners' paragraph writing performance. An explanatory sequential mixed-methods research approach with a pre-test post-test quasi-experimental design was employed. Two intact sections were selected using convenience sampling, and they were randomly assigned to experimental and control groups. The experimental group used ProWritingAid to receive automated feedback on their paragraph drafts, whereas the control group received conventional teacher feedback. Data were collected using paragraph writing tests and interviews. The quantitative data were analyzed using paired-samples t-tests and one-way MANOVA. On the other hand, a systematic data analysis procedure was employed to analyze the qualitative data. The findings revealed that students' writing performance was significantly improved, especially in grammar, vocabulary, mechanics, and overall writing performance. However, there was no considerable difference in content and organization. Additionally, the experimental group significantly outperformed the control group in grammar, vocabulary, and mechanics with large effect sizes (η^2) of .182, .201, and .236, respectively. Furthermore, students perceived using the automated tool positively. They appreciated the improved writing performance, ease of use, immediate feedback, consistent feedback, and increased motivation. Therefore, we concluded that utilizing automated feedback tools is a practical approach that can enhance students' writing performance in the Ethiopian higher education EFL context although teacher feedback remains important for higher-order aspects of writing. Overall, it is observed that automated written feedback can be an alternative approach to support university EFL learners' writing development.

Keywords: automated written corrective feedback, ProWritingAid, writing aspects, writing performance

[1]Corresponding author's Email:manaye.mesfin@kiot.edu.et

Article information:

Received: 10Sept, 2025 Revised: 31Jan, 2026
Accepted:19Feb, 2026 Available online: 31March, 2026

Doi: http://doi.org/10.20372/ajids.2026.2568

1. Introduction

Writing is one of the most significant language skills, and EFL students demand it for better academic performance. In educational settings, it enables students to use the language creatively and extend their learning (Fareed*et al*., 2016; Tseng, 2019). Writing helps students organize ideas, construct arguments, and convey meanings as it supports linguistic, cognitive, and metacognitive abilities (Klimova, 2013). It also improves learners' workplace communicative competencies, including report, letter, and resume writing (Bora, 2023). Furthermore, it facilitates learners' social interactions and helps students increase their cognizance of various aspects of life (Andrei *et al*., 2019).

However, both native and non-native learners find writing in English one of the most challenging skills. First, students face difficulties in generating, organizing, and translating ideas into readable texts (Jahin&Idrees, 2012). Second, unlike other language skills, writing requires a high degree of accuracy as it is a recorded form of communication that must be understood without cues (Hyland, 2003; Harmer, 2004). Finally, learning to write is difficult because of its multistage nature, including brainstorming, drafting, and editing (Rao, 2017; Bora, 2023). These challenges in the teaching of writing indicate the need for effective instructional support in writing classes.

Similarly, producing clear and engaging written text is a major challenge for Ethiopian EFL students at all levels (Geremew, 1999; Italo, 1999; Amlaku, 2010). Even after completing secondary school education, students showserious deficiencies in writing. They struggle not only to write continuous texts but also to construct a meaningful sentence at the tertiary level (Ebabu, 2018; Dejen&Seime, 2023). Overall, these findings indicate that despite ongoing research efforts, EFL students in Ethiopia continue to face significant challenges in developing their writing skills. Given these persistent challenges, applying effective pedagogical interventions

such as written corrective feedback (WCF) has become crucial for supporting students' writing development.

Hence, in teaching writing, WCF is one of the fundamental pedagogical interventions that enable students to restructure and develop their ideas. Students expect to receive comments on their writing because they believe that error correction improves writing (Hyland, 2003). Similarly, feedback enhances writing, and it is a key factor in determining writing performance in second language instruction (Ellis, 2009); it improves students' writing quality by 16% (Graham & Harris, 2016). Consequently, WCF is widely regarded as an inextricable component of effective writing instruction.

Although teacher WCF is essential to teaching writing, it has some shortcomings. It is a time-consuming and arduous task for teachers (Attali*et al*., 2010). Teachers might be unable to respond to each student's work individually, quickly, or in a way relevant to improve the content(Grimes&Warschauer, 2010; Lee *et al*., 2009). WCF is also exposed to variation due to teachers' individual differences, such as attitude and motivation (Goldstein, 2006; Ferris *et al*., 2012). Moreover, the problem worsens when class size is large (Hyland & Hyland, 2019). These limitations highlight the need to explore alternative feedback mechanisms.

In response to the weaknesses of the conventional teacher-led written corrective feedback (CTWCF), automated written corrective feedback (AWCF) has emerged. When compared to CTWCF, it results in greater improvement on students' writing performance (Goldstein, 2006; Hartshorn, 2010; Ferris *et al*., 2012). By generating consistent, immediate, and impartial feedback, AWCF can address the limitations of CTWCF (Fang, 2010; El Ebyary&Windeatt, 2010). However, it has limitations (Aluthman, 2016; Lai, 2010). It could be too complex for ESL and EFL learners at a lower proficiency level (Aluthman, 2016), and some automated systems provide repetitive feedback that may confusestudents (Lai, 2010).

In addition to the empirical evidenceon its effectiveness, AWCF has theoretical underpinnings, including the Noticing Hypothesis (Schmidt, 1990), the Interaction Hypothesis (Long, 1996), the Output Hypothesis (Swain, 1985), and Computer-Assisted Language Learning (Chapelle, 2001).

Learners can notice the gaps between their interlanguage and target forms (Schmidt, 1990), and automated feedback helps them notice and correct errors. Learners interact with computer applications, and according to Long (1996), interaction is central to language development. Furthermore, feedback pushes learners to produce accurate and appropriate language (Swain, 1985). Finally, computers provide personalized, immediate, and consistent feedback that supports language learning (Chapelle, 2001).

Considering the importance of AWCF in writing instructions, a range of AWCF systems with sophisticated language processing technologies have been developed. Such systems allow students to access alternative sources of WCF (Zhang, 2017). For example, Criterion, My Access,Pigai, Grammarly, Cambridge Write & Improve, and ProWritingAid (PWA) are among others, and they are used to score and edit writing (Woodworth &Barkaoui, 2020).

In this study, PWA was employed for three main reasons. First, it is one of the most recently developed AWCF tools and has evolved into an advanced automated tool with user-friendly features (ProWritingAid, n.d.). Second, most AWCF studies to date were conducted on Pigai, Criterion, and Grammarly. A systematic review of 83 Social Sciences Citation Index (SSCI) indexed journal articles on 31 different types of AWCF tools indicated that these tools were widely covered (Shi &Aryadoust, 2024). PWA,on the contrary, was used rarely, indicating a literature gap. Finally, unlike other AWCF tools, it can be used without a subscription fee with a single account, making it a practical option in educational research contexts (Ariyanto*et al*., 2019).

To determine the relevance of AWCF toenhance students' writing performance in Ethiopian higher education, we conducted a preliminary investigation through document analysis and interviews. According to the findings from the document analysis, there were frequent errors, including grammar, vocabulary, spelling, capitalization, punctuation, content, and organization in students' texts. In addition, teachers' practices of WCF provision were inconsistent: providing no feedback, putting marks for grading, underlining errors, and writing general comments.

Additionally, interview participants perceived that WCF helps students improve their writing performance. Consistently, Leong and Lee (2018) found that WCF assists students in identifying gaps in their writing. However, there are limitations in WCF delivery practices. First, some teachers are reluctant to provide feedback on writing assignments. Second, most students do not have opportunities to see their errors after submitting written assignments. Third, AWCF tools are rarely utilized in writing classes. These findings show that one possible reason for the persistent writing problems may be the nature and quality of corrective feedback. Exploring alternative WCF mechanisms is thus determinant for improving students' writing performance.

The problems we have observed in feedback provision practice led us to search for global studies on AWCF using PWA. A few studies examined PWA and students' writing performance and found positive results (Nasution& Fatimah, 2018; Ariyanto*et al*., 2019; Pitukwong&Saraiwang, 2024). It helps EFL learners notice errors, improve writing performance, and increase motivation (Nasution& Fatimah, 2018). Using PWA, students showed significant improvement in grammar, organization, mechanics, vocabulary, and unity and coherence (Pitukwong&Saraiwang, 2024). Furthermore, it reduces teachers' workload. Therefore, both students and teachers perceived AWCF supported by PWA positively (Ariyanto*et al*., 2019; Fitria, 2023).

Despite the contributions of previous studies, the present study seeks to address several gaps. Nasution and Fatimah (2018) examined the use of PWA among high school students; however, this study focuses on undergraduate EFL learners. Methodologically, Ariyanto*et al*. (2019) employed a survey design to examine students' and teachers' perceptions; the present study utilizes a quasi-experimental design, focusing exclusively on students. Pitukwong and Saraiwang (2024) compared the effectiveness of PWA and Writing Punch, whereas this study contrasts AWCF supported by PWA with CTWCF. Moreover, while Fitria (2023) examined the grammar and spelling-checking role of PWA, the present study investigates its effects on the grammar, vocabulary, mechanics, content, and organization aspects of writing.

In Ethiopia, Italo (1999), Assres (2021), Tibebu*et al*. (2022), and Baymot*et al*. (2023; 2024) investigated different WCF types, but they focused on human-generated feedback. More recently, a few local studies, for instance, Sime *et al*. (2024) and Demissie*et al*. (2025),have

examined the effects of technology-mediated writing instructions and established their pedagogical values. However, the feedback remains human-generated. This suggests that studies examining the effects of AWCF on Ethiopian EFL learners' writing performance remain limited. Hence, the following research questions were designed.

1) Does AWCF improve undergraduate EFL learners' paragraph writing performance?
2) Is there any significant difference between the effect of AWCF and CTWCF on students' paragraph writing performance?
3) How do students perceive learning to write through AWCF scaffolding?

2. Materials and Methods

Having followed a pragmatic worldview, the study employed a mixed-methods research approach. It incorporates both quantitative and qualitative data within a single study (Creswell & Creswell, 2018). In particular, the explanatory sequential (QUAN-qual) approach was used because quantitative data collection and analysis were conducted before qualitative data collection and analysis. Since there was no random selection of study participants, the study employed a pre-test post-test control group quasi-experimental design.

2.1. Participants of the Study

The study was conducted at Wollo University, Kombolcha Institute of Technology (KIoT), and the participants were first-year undergraduate students of the campus. This is due to the reason that Communicative English Language Skills courses are offered in the first-year of university education. Among the 30 first-year sections enrolled at KIoT in the 2024/25 academic year, two intact sections were selected using convenience sampling. Because they were taught by one of the authors, the sections were appropriate to run the experiment smoothly. Once the homogeneity of the two sections was established using pre-test scores, they were randomly assigned to experimental and comparison groups. The experimental and comparison groups consisted of 46 and 45 students,respectively. In sum, 91 first-year undergraduate students of KIoTwere participants of the study. In addition, six interviewees (three male and three female) were selected from the experimental group using purposive sampling.

2.2. Data Collection Instruments

In this study, both quantitative and qualitative data were collected. The quantitative data were gathered using paragraph writing pre- and post-tests. The pre-test was administered before the intervention for two main purposes: to determine students' writing performance at the outset and to check the homogeneity of students in the two groups. Again, after the completion of the intervention, a post-test was administered with the aim of examining the effects of AWCF on students' writing performance. Efforts were made to match the difficulty of the writing prompts across the two testing conditions while varying the topics to avoid practice effects.

On the other hand, the qualitative data were collected using semi-structured interviews. The interview for each participant consisted of questions related to students' experience of using AWCF, how AWCF helps them, comparison of AWCF and CTWCF, confusion and difficulty of feedback, and motivation to revise and write. Interviews were conducted on March 30/2025, at the English language improvement center (ELIC) office. We ensured that all interview participants wereable to communicate in Amharic effectively. Thus, to avoid potential communicative difficulties, Amharic was considered appropriate for the interview. Before the interview began, the purpose of the interview and the voluntary nature of participation were explained. Then, participants were contacted individually. During the interview, there was an audio recording using a smartphone to maintain credibility. Then, the data was transcribed verbatim and translated into English.

2.3. Validity and Reliability of Instruments

Validity is the degree to which an instrument accurately measures what it is intended to measure (Creswell & Creswell, 2018). In this study, we ensured the validity of the paragraph writing tests by obtaining comments from three experienced English language instructors. Additionally, 20 first-year undergraduate students from randomly selected sections were pilot-tested. Although both the pilot and main studies were conducted at KIoT, participants were selected from different cohorts of first-year undergraduate students: they werethe 2023/24 and 2024/25 batches, respectively. The result of the pilot study revealed that the test was valid to measure students' writing performance based on predetermined criteria of grammar, vocabulary, mechanics, organization, and content.

Two experienced university English language instructors scored the tests to minimize researcher bias. One of the authors trained the two raters to apply the analytical scoring rubric developed by Agan and Deniz (2019) to score the writing tests. The training lasted four hours. Then, they independently scored students' paragraphs considering grammar, vocabulary, mechanics, organization, and content aspects of writing. To check the reliability of paragraph writing tests, interrater reliability was calculated, and the result was found to be excellent as the Intraclass correlation coefficient (ICC) of the two raters was > 0.90. Finally, the average scores of the two raters were analyzed.

2.4. Data Collection Procedures

Data were collected in three stages. First, a paragraph writing pre-test was administered to both the experimental and control groups before the intervention. In this test, students in both groups were required to write an expository paragraph of 120 words using one of the three alternative topics given. Then, the result of the pre-test was analyzed to ensure homogeneity of the two groups. Second, following the completion of the intervention, a paragraph writing post-test was administered with similar but not the same paragraph writing prompts. Finally, interviews were conducted with selected students from the experimental group.

2.5. The Intervention Procedures

To run the experiment, an intervention material was prepared by adapting the paragraph writing lessons from the module for 'Communicative English Language Skills I' (FLEn 1011). The material consisted of paragraph writing lessons, paragraph writing activities, and training on how to use PWA. All paragraph writing lessons and activities were combined accordingly because they were presented in a special writing class once a week over continuous periods. Therefore, the intervention material helped us implement the treatment effectively. The experiment took a total of nine weeks, classified into three phases: the pre-intervention, the intervention, and the post intervention phases.

The pre-intervention phase (two weeks)

In the pre-intervention stage, administration of the pre-test and introduction of the main concepts of paragraph writing to both the treatment and comparison groups were performed. Additionally, only the treatment group was given training on how to use PWA for three hours.

The intervention phase (six weeks)

This stage constituted the main experiment. Students in both groups engaged in similar paragraph writing activities but under different feedback conditions. Each student wrote six expository paragraphs (one per week)using the prompts adapted from the teaching material for FLEn 1011 (see Appendix I). For each writing activity, students in the treatment group received AWCF supported by PWA using internet-connected computers in a computer laboratory. On the other hand, students in the control group received CTWCF from their English language teacher.

The post-intervention phase (the final week)

This was the last phase of the intervention. At this stage, a paragraph writing post-test was administered to students in both the treatment and comparison groups. Since it was an immediate post-test, it was administered one week after the completion of the intervention.

2.6. Methods of Data Analysis

To analyze the quantitative data, descriptive and inferential statistics were employed. Students' paragraph writing tests were scored analytically by assigning 20 points for each aspect of writing (i.e., grammar, vocabulary, mechanics, content, and organization). Then, mean and standard deviation were calculated to determine the difference between the pre- and post-test scores of participants. In addition, paired-samples t-tests were computed to examine whether there were statistically significant differences between the pre-test and post-test scores of the experimental group. Finally, a one-way MANOVA was conducted to assess whether post-test scores differed significantly between the experimental and control groups.

Before running the paired samples t-tests and the one-way MANOVA, assumption tests were run using SPSS version-27 computer software. The results of the Kolmogorov-Smirnov and Shapiro-Wilk tests were > .05. This revealed that the data were approximately normally distributed. To check homogeneity, Levene's test of homogeneity of variances was computed, and the p-value (>.05) showed that the assumption of homogeneity was met. In addition to these tests,

multivariate and univariate normality, Linearity, and Box's test of equality of covariance matrices were tested, and all the assumptions were met. Therefore, it was possible to run the parametric tests.

On the other hand, the qualitative data from students' interviews were analyzed inductively using Creswell and Creswell's (2023) seven-step procedure, involving transcription, coding, theme development, and member checking to ensure credibility. First, all recorded interview data were transcribed verbatim and interpreted into English. Second, we read through the transcripts repeatedly to gain a comprehensive understanding of the participants' views. Third, descriptive codes were assigned to the data. Fourth, we grouped the data into different themes. Fifth, conceptually related categories were interconnected. Sixth, themes were presented using analytical narratives followed by appropriate quotations. Finally, the credibility of the data was validated using member checking.

2.7. Ethical Considerations

For the purpose of protecting study participants, various ethical standards were adhered. First, participants completed a written consent form that contained the purpose of the experiment, the intervention procedure, and voluntary participation. They were also informed that they could withdraw from the intervention at any time. Second, the confidentiality of the data from participants was declared. Third, all the data used in the study were collected from participants. Fourth, to uphold the principle of justice, we ensured that students had equal opportunities to contribute to and benefit from the study. Finally, students in the control group were given a makeup class on how to use PWAfor four hours as compensation. To prevent data contamination, we arranged the compensation session after the completion of the intervention.

3. Results and Discussion

3.1.Results of the Pre-test

Table 1. Levene's Test of Equality of Variances

Writing Aspects	Grammar	Vocabulary	Mechanics	Organization	Content	Overall
F	.010	.392	.002	1.133	.000	.030
Sig	.922	.533	.966	.290	.994	.862

The result in Table 1 revealed no statistically significant difference between the two groups, considering the five main aspects of writing and overall writing performance. As shown in Table 1, the p-valuesof Levene's test of equality of variances for all writing aspects and overall writing performance are >.05. Hence, the two groups were homogeneous in their paragraph writing performance before the intervention.

3.2.AWCF and Students' Paragraph Writing Performance

Table 2.Pre-test and Post-test Scores of the Experimental Group

Pairs	Writing Aspects	Tests	N	Mean	Sd.	Mean df.	T	Df	Sig. (2-tailed)
1	Grammar	Pre-test	46	9.13	1.96	-4.13	-19.46	45	.000
		Post-test	46	13.26	2.43				
2	Vocabulary	Pre-test	46	8.28	1.83	-3.59	-17.09	45	.000
		Post-test	46	11.87	2.16				
3	Mechanics	Pre-test	46	9.69	1.80	-3.35	-18.98	45	.000
		Post-test	46	13.04	2.29				
4	Content	Pre-test	46	8.91	2.09	.065	0.465	45	.645
		Post-test	46	8.85	2.22				

5	Organization	Pre-test	46	8.80	2.23	-.456	-2.84	45	.007
		Post-test	46	9.26	2.13				
6	Total	Pre-test	46	44.83	9.09	-11.45	-27.05	45	.000
		Post-test	46	56.28	9.77				

As Table 2 indicates, the post-test results confirmed that there were statistically substantial improvements in grammar (M=4.13; p<.001; t= 19.46), vocabulary (M=3.59; p<.001; t= 17.09). Similarly, there were considerable improvements in mechanics (M=3.35; p < .001; t=18.98), organization (M=.456; p=.007; t=2.28). Finally, the mean difference of 11.5 and p<.001 revealed that students' overall writing performance was significantly improved. On the other hand, p = .645; t = .465indicated no significant difference in content development. This implies that the AWCF intervention meaningfully improved all aspects of writing except content.

3.3. Effects of AWCF and CTWCF on students' paragraph writing performance

Table 3.Multivariate Test

	Value	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Pillai's trace	.688	37.540a	5.000	85.000	.000	.688
Wilks' lambda	.312	37.540a	5.000	85.000	.000	.688
Hotelling's trace	2.208	37.540a	5.000	85.000	.000	.688
Roy's largest root	2.208	37.540a	5.000	85.000	.000	.688

The results in Table 3 revealed that AWCF and CTWCF had statistically significant multivariate effects on students' writing post-test scores. Wilks' Lambda = .312, F (5, 85) = 37.54, p < .001, with a large effect size (η^2 = .688). This indicates that feedback (AWCF and CTWCF) had

significant effects on students' writing performance. However, the multivariate test could not detect differences between the AWCF and CTWCF. Therefore, to further identify the difference between the two feedback conditions, the test of between-subjects effects was calculated.

Table 4.Test of Between-Subjects Effects

	Grammar	**Vocabulary**	**Mechanics**	**Content**	**Organization**	**Total**
F-Value	19.767	22.424	27.555	4.872	0.406	9.471
Sig. (p-value)	.000	.000	.000	.030	.526	.003
Partial Eta Squared (η^2)	.182	.201	.236	.052	.005	.096

As the results in Table 4 indicate, there were statistically significant differences between the experimental and control groups in grammar ($p < .001$), vocabulary ($p < .001$), mechanics ($p < .001$), content ($p = .030$), and total writing score ($p = .003$). However, regarding organization ($p = .526$), no significant differences were observed. The Partial Eta Squared (η^2) showed large effect sizes for grammar (.182), vocabulary (.201), and mechanics (.236). However, the effect sizes were medium (.096) for overall writing performance and small (.052) effect sizes for content, and the feedback conditions had almost no effect (η^2 .005) on paragraph organization.

Although the test of between-subjects effects indicated the existence of significant differences between the experimental and control groups, it cannot give any clue about the group that outperformed in the paragraph writing post-test. Hence, we interpreted the estimated marginal means and 95% confidence intervals to identify which group outperformed.

Table 5. Estimated Marginal Means and 95% Confidence Intervals of the Post-test

Variables	Group	Mean	Std Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Grammar	Experimental	13.26	.340	12.59	13.94
	Control	11.11	.344	10.43	11.79

Vocabulary	Experimental	11.87	.340	11.19	12.55
	Control	9.58	.344	8.90	10.26
Mechanics	Experimental	13.04	.339	12.37	13.72
	Control	10.51	.343	9.83	11.20
Content	Experimental	8.85	.332	8.19	9.51
	Control	9.89	.335	9.22	10.56
Organization	Experimental	9.26	.313	8.64	9.88
	Control	8.98	.316	8.35	9.61
Total	Experimental	56.28	1.42	53.46	59.10
	Control	50.07	1.44	47.21	52.92

As we can see from Table 5, the experimental group significantly outperformed the control group in grammar, vocabulary, mechanics, organization, and overall writing performance. The experimental group achieved a higher estimated mean score in grammar (M = 13.26; SE = 0.34), vocabulary (M=11.87; SE = 0.34), mechanics (M = 13.04; SE = 0.34), and organization (M = 9.26; SE = 0.313). In general, the experimental group's total score (M = 56.28; SE = 1.42) is higher than the control group's (M = 50.07; SE = 1.44). Therefore, there were statistically significant difference in the writing performance of the experimental and control groups, favoring the AWCF group.

In addition, the 95% confidence intervals for the two groups do not overlap in grammar, vocabulary, mechanics, and in the total scores. No overlap in the 95% confidence interval revealed that there were potentially considerable differences between the two groups. Exceptionally, the control group had a slightly higher mean score (M=9.89) compared to the AWCF group (M=8.85) in content. In addition, there are slight overlaps in the 95% confidence intervals. Hence, the difference was not statistically significant.

3.4. Students' Perceptions of Learning to Write through AWCF Scaffolding

Having gained insights about the positive effects of AWCF from the analysis of test results in the two testing conditions, students were interviewed to explore their perceptions of learning to write through AWCF scaffolding, categorizing them into five main themes. These are: students' experience of using AWCF, how AWCF helps them improve their writing, confusion and difficulty of feedback, comparison of AWCF and CTWCF, and motivation to revise and confidence in writing.

Students' experience of using AWCF

According to the interview data, most students had no experience of receiving AWCF. However, a few reported limited exposures to AI tools such as ChatGPT. One of the participants explained, "I frequently use ChatGPT to write different letters for scholarships because I did not have confidence in my writing skills. It benefited me a lot to achieve grammatical accuracy and use punctuation marks and capital letters appropriately." (S1)

How does AWCF helpstudents improve their writing?

Regarding the importance of PWA, the interview data indicated that most participants considered it highly beneficial. They explained that PWA is a user-friendly, freely accessible feedback tool that provides feedback on grammar, spelling, punctuation, sentence length, and vocabulary. Participants also confirmed that PWA provided indirect feedback via color-coded lines for different types of errors and explanations of alternative correct forms. Furthermore, they appreciated the immediacy of feedback that can be provided anytime and anywhere. Let us see an excerpt from one of the participants.

> I found PWA user-friendly, and the way it explained my errors helped me identify different grammatical concepts, including how to use articles and prepositions, the main rules of subject-verb agreement, capitalization, punctuation, and spelling rules. In addition, the immediate feedback with no subscription fee, provision of indirect feedback by underlining the errors, and so forth make PWA a helpful writing assistant. (S3)

Confusion and difficulty

Most of the participants demonstrated that the feedback they received from PWA was mostly easy to understand, and there was no confusion in identifying errors. However, a few participants reflected that some explanations were unclear. As one participant noted, "Some grammar explanations were difficult to understand, and I was confused by some technical terms produced by PWA." (S6)

Comparison of AWCF and CTWCF

When asked to compare AWCF and CTWCF, most participants reflected that they preferred AWCF to CTWCF because AWCF can generate detailed and immediate feedback, while teachers often failed to provide timely WCF. Moreover, some students explained that some teachers did not give any feedback on their writing assignments. However, a few students preferred CTWCF as they were demotivated because of unexpectedly low scores generated by PWA. They also noticed that the feedback was limited to grammar, vocabulary, punctuation, and other surface-level aspects of writing, with almost no attention to content. The following quote is taken from one of the participants.

> PWA supported me to revise my paragraphs by generating detailed feedback on many aspects of writing, including grammar, vocabulary, punctuation, and style, but the contents of my paragraphs were not commented on by the tool." Therefore, it would be more effective if the AWCF were supported by teacher feedback. (S5)

Motivation to revise and confidence in writing

Most participants confirmed that PWA helped them develop motivation to revise their paragraphs. In addition, students explained that they thought they could not write a paragraph because it was a difficult task, but PWA developed their confidence in writing. However, there are a few students who were demotivated by excessive correction of errors. As one student explained, "Over correction and low scores in different aspects of writing demotivated me to use PWAthough I have changed this feeling through repeated practice of using the website." (S6)

4. Discussion

The first research question aimed to determine whether AWCF improved students' paragraph writing performance, and the findings revealed that AWCF, using PWA, enhanced students'

paragraph writing. The post-test scores of the experimental group revealed a statistically significant improvement in students' paragraph writing, particularly in grammar accuracy, word choice, lexical range, mechanics, and overall paragraph writing skills.

The findings of this study aligned with some previous studies. For instance, Soleimani and Moqimi (2024) found that PWA is among the most important online applications for the enhancement of EFL/ESL writing. In addition, AWCF tools such as Grammarly and PWA help students improve writing, especially linguistic aspects (Wahyuda*et al*., 2022; Damayanti &Santosa, 2024).Similarly, Wang (2013) confirmed that automated feedback increases grammatical accuracy and learners' autonomy. On the other hand, there are studies which claimed the limited effectiveness of AWCF. For instance, Aluthman (2016) found that sometimes AWCF could be too complex for learners at a lower proficiency level.

The second research question was designed to determine whether there is a significant difference between the effects of AWCF and CTWCF on students' writing performance across different aspects of writing and students' overall writing performance. The findings from a one-way MANOVA confirmed that there was a statistically significant enhancement of students' grammar, vocabulary, mechanics, and overall writing performance in both the AWCF and CTWCF groups. However, the experimental group significantly outperformed the control group. However, there was no practical difference between the two groups in content and organization.

Consistently, Wahyuda et al. (2022) reported that grammar and mechanics aspects of Indonesian students were significantly improved because of AWCF. However, content-related issues required teacher mediation. Similarly, Perdana and Farida (2019) confirmed the effectiveness of online feedback tools to correct grammar mistakes. However, the findings of this study do not align with Lai's (2010), as he found that the feedback from AWCF tools is formulaic and causes confusion.

Finally, the third research question aimed to explore students' perceptions of learning to write through AWCF scaffolding. The findings of the interviews indicated that although most of the students did not have experience of using PWA, they perceived it positively. They believed that

PWA helped them improve their writing because it provides them with immediate and consistent feedback on different writing aspects. Students also perceived that PWA is user-friendly software. Since it can help them easily identify and immediately correct errors, their motivation to write and confidence in their writing were considerably increased.

Previous studies correspondingly confirm the positive views of students on AWCF. For instance, Barrot (2021) confirmed that students who were exposed to AWCF appreciated the metalinguistic explanations and indirect feedback using color-coded visual cues because the explanations and colored lines helped them identify and correct errors in their writing. Similarly, Nasution and Fatimah (2018) and Aboulfotoh (2024) reported that students valued the feedback from AWCF tools because there is no fear of negative judgment. Moreover, AWCF enhances students' motivation and confidence to write. Even so, there are studies showing negative views of students towards AWCF. For example, Sherafati *et al*. (2020) documented that students do not appreciate the role of computers as a feedback source, but accept AWCF as a supplement to teacher feedback.

5. Conclusions

The main purpose of this study was to examine the effects of AWCF on first-year undergraduate students' paragraph writing performance and their perception of learning to write through its scaffolding. The findings revealed that AWCF significantly improved students' paragraph writing performance. In particular, grammar accuracy, vocabulary, mechanics, and total scores of students were practically improved. However, AWCF resulted in a limited impact on the organization of ideas and no impact on content development.

The results also showed that students exposed to AWCF using PWA significantly outperformed those in the CTWCF group, but no substantial difference was observed between the two groups in idea organization. Furthermore, the control group outperformed the experimental group in content development although the difference was not substantial. Therefore, AWCF has practical effects on grammar, vocabulary, mechanics, and overall writing performance, but it did not have significant impacts on organization and content, emphasizing that teacher feedback remains important for holistic writing development.

Finally, students perceived AWCF supported by PWA positively. They appreciated the immediate and consistent feedback, the detailed explanations, and the alternative corrections provided by PWA. The tool enhanced their motivation and confidence in writing. Nonetheless, some limitations, including overcorrection, vague explanations, and a major focus on surface-level aspects were identified.

6. Limitations and Future Directions

Although all the instruments were validated, the study is not free from limitations. First, only forty-six undergraduate EFL students participated in the experiment, and this may affect the generalizability of the findings. Second, the experiment was made taking only one AWCF tool, while we had justifications. Third, since there was a time limitation, the study did not consider the retention effects of AWCF on students' paragraph writing performance.

Given the study's limitations, further efforts are needed to examine the effectiveness of AWCF. Because of the development of AI, there exist diverse automated tools and computer applications that can assist with classroom writing instructions. Therefore, future researchers are recommended to conduct longitudinal studies with a relatively large sample size and examine the retention effects of AWCF on students' paragraph writing performance. Additionally, to determine the relative effectiveness of different digital feedback tools, future researchers should study different AWCF software in the Ethiopian ELT context.

Acknowledgements

First, we would like to thank Wollo University for funding this study. Our sincere appreciation is also extended to first-year undergraduate students at Wollo University, Kombolcha Institute of Technology in the 2024/2025 academic year for their willingness to participate in this study.

Competing interest

The authors reported no competing interests.

Authors' Contribution

The corresponding author conducted the study, getting all the necessary permissions from Wollo University to run the experiment. Then, data collection, data analysis, report writing, and

proofreading were done by both the corresponding author and the co-author. In addition, the co-author, besides conceptualizing the research issue, guided the corresponding author in deciding on the most appropriate research methodology, writing the report, and providing all necessary comments and guidance.

References

Aboulfotoh, D. M. A. (2024). Applying ProWritingAid and BASDELL model in Business English writing classes for management sciences Students. *Journal of Research in Curriculum Instruction and Educational Technology*, *10*(1), 135-180.

Agan, Ş. İ., &Deniz, S. (2019). A rubric study for assessing paragraph level written texts. *Journal of Education and Training Studies*, *8*(1), 42-50.https:// doi.org/ 10.11114/ jets. v8i1.4610

Aluthman, E. S. (2016). The effect of using automated essay evaluation on ESL undergraduate students' writing skill. *International Journal of English Linguistics*, *6*(5),54-67.http://dx.doi.org/10.5539/ijel.v6n5p54

Amlaku, B. (2010, March 19–21). Language policies and the role of English in Ethiopia. Paper presented at the 23rd annual IATEFL BESIG conference, Bielefeld, Germany.

Andrei, E., Ellerbe, M., & Kidd, B. (2019). "What am I going to do?" A veteran teacher's journey of teaching writing to newcomer English language learners. *TESOL Journal*, *10*(2), 1-19. https://doi.org/10.1002/tesj.413

Ariyanto, M. S. A., Mukminatien, N., &Tresnadewi, S. (2019). Students' and teachers' perceptions towards the implementation of ProWritingAid and teacher feedback.PhD. dissertation, State University of Malang.

Attali, Y., Bridgeman, B., & Trapani, C. (2010). Performance of a generic approach in automated essay scoring. *The Journal of Technology, Learning and Assessment*, *10*(3), 1-16.

Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, *36*(4), 584-607.https:// doi.org/ 10.1080/09588221.2021.1936071

Bekele, T., Seyoum, G., &Tibebu, A. (2022). Comparing the Effects of Direct and Indirect Feedback on Students' Writing Performance. *East African Journal of Social Sciences and Humanities, 7*(1), 33-46.

Bora, P. (2023). Importance of writing skill to develop students' communication skill. *Journal for Research Scholars and Professionals of English Language Teaching, 7*(35), 1-6.https://doi.org/10.54850/jrspelt.7.35.009

Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge University Press.

Creswell, J.W., &Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches (6th ed.).* Sage Publications.

Creswell, J.W.,& Creswell, J.D. (2018).*Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications.

Damayanti, K. D., &Santosa, M. H. (2024). The effect of ProWritingAid writing tool on the writing skills of English as a foreign language (EFL) students: A systematic literature review. *Indonesian Journal of Multidisciplinary Research, 4*(2), 469-478.http://dx.doi.org/10.17509/xxxx.xxi

Dejene, M. &Seime, K. (2023). Investigating the relationship between university students' perceptions of themselves as writers and their writing performance Worldwide: *Journal of English Language Teaching, 10*(1), 75-89.

Demissie, G., Amogne, D., &Gashaw, A. (2025). The effect of using Google Docs on Ethiopian EFL students' collaborative writing improvement. *Australian Journal of Applied Linguistics, 8(*3).https://doi.org/10.29140/ajal.v8n3.102622

Ebabu, T. (2018). A study on English as a foreign language writing: Instruction, nature of tasks, and performance of students: A case of North Shoa Preparatory Schools. *Ethiop.j.soc.lang.stud., 5* (2), 61-92.

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.

Ellis, R. (2009). Corrective feedback and teacher development.*L2 Journal, 1*(1), 3–18. https://doi.org/10.5070/L2.V1I1.9054

Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *Journal of Educational Technology & Society*, *13*(3), 246-256.

Fareed, M., Ashraf, A., & Bilal, M. (2016). ESL learners' writing skills: Problems, factors and suggestions. *Journal of Education and Social Sciences, 4*(2), 81-92.https:// doi.org/ 10.20547/jess0421604201

Ferris, D. R., Liu, H., Sinha, A., &Senna, M. (2012). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing, 22*(3), 307–329.https:// doi.org /10.1016 /j.jslw.2012.09.009

Fitria, T. N. (2023). ProWritingAid as ai-powered writing tool: The performance in checking grammar and spelling of students' writing. *Polingua (Scientific Journal of Linguistics, Literature and Language Education), 28*, 65-75.

Geremew, L. (1999). A Study of the requirements in writing for academic purposes at AAU.AAU, Unpublished Ph. D. dissertation.

Goldstein, L. (2006). Feedback and revision in second language writing: Contextual, teacher, and student variables. In K. Hyland & F. Hyland (Eds.*), Feedback in second language writing: Contextual and issues* (pp. 185–205). Cambridge University Press.

Graham, S., & Harris, K. (2016). Feedback and improvement in student writing.Learning and Instruction, 45, 23–34.

Grimes, D., &Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *Journal of Technology, Language, and Assessment, 8*(6), 1-43.

Harmer, J. (2004). *How to Teach Writing*. Longman

Hartshorn, K. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy*?TESOL Quarterly, 44*(1), 84-109. https://doi.org/10.5054/tq.2010.213781

Hyland, K. (2003). *Second Language Writing*.Cambridge University Press.

Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 Writing.In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 1–22).Cambridge University Press.

Italo, B. (1999). A Comparison of the effectiveness of teacher versus peer feedback on AAU students' writing revisions.AAU, Unpublished Ph. D. dissertation.

Jahin, M., &Idrees, M. (2012). Challenges in EFL writing.*English Language Teaching, 5*(3), 33–40.

Klimova, B. F. (2012). The importance of writing.*Indian Journal of Research*, *2*(1), 9-11.

Lai, Y. H. (2010). Which do students prefer to evaluate their essays: Peers or a computer program? *British Journal of Educational Technology*, *41*(3), 432-454.https://doi.org/10.1111/j.1467-8535.2009.00959.x

Lee, C., Wong, K. C. K., Cheung, W. K., & Lee, F. S. L. (2009). Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation. *Computer Assisted Language Learning*, *22*(1), 57-72.https://doi.org/10.1080/09588220802613807

Leong, L., & Lee, F. (2018). Teacher and student perceptions of corrective feedback.*ELT Journal*, *72*(2), 148–159.

Long, M.H. (1996). The role of linguistic environment in second language acquisition.In W.C. Ritchie, & T.K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). CA: Academic Press.

Nasution, A., & Fatimah, S. (2018). The use of ProWritingAid web in editing students' writing. *Journal of English Language Teaching*, *7*(2), 362-368.

Nigus, A. (2021). EFL writing teachers' and students' perceptions, preferences, and practices of written corrective feedback techniques in improving grammatical accuracy. Abyssinia *Journal of Business and Social Sciences*, *6*(1), 24-35.

Perdana, I., & Farida, M. (2019). Online grammar checkers and their use for EFL writing. *Journal of English Teaching, Applied Linguistics and Literatures (JETALL)*, *2*(2), 67-76.

Pitukwong, K., &Saraiwang, S. (2024). Exploring the effectiveness of digital writing tools on Thai EFL students' writing. Contemporary Educational Technology, 16(3).https:// doi.org/10.30935/cedtech/14808

ProWritingAid. (n.d.). ProWritingAid: AI writing assistant. Retrieved from https://prowritingaid .com

Rao, P. S. (2017). Developing writing skills among the EFL/ESL learners. *Research Journal of English, 2*(3), 52-63.

Schmidt, R. (1990). The role of consciousness in second language learning.*Applied Linguistics, 11*(2), 129-58.https://doi.org/10.1093/applin/11.2.129

Sherafati, N., Largani, F. M., &Amini, S. (2020). Exploring the effect of computer-mediated teacher feedback on the writing achievement of Iranian EFL learners: Does motivation count? *Education and Information Technologies, 25*(5), 4591-4613. https:// doi.org/ 10.1007/s10639-020-10177-5

Shi, H., &Aryadoust, V. (2024).A systematic review of AI-based automated written feedback research. *ReCALL, 36*(2), 187-209.https://doi.org/10.1017/S0958344023000265

Sime,T., Gencha M. &Olamo, T. (2024). Effects of blended instruction on students' paragraph writing performances: the case of first year Health science pharmacy students at Pharma College Hawassa Campus, Ethiopia, *Cogent Education, 11*(1).https:// doi.org /10.1080 /2331186X.2024.2321309

Soleimani, H., &Moqimi, J. (2024). Teaching ESL/EFL writing skill website evaluation: the Purdue Online Writing Lab and ProWritingAid. *Interdisciplinary Studies in English Language Teaching, 2*(1), 100-114.https://doi.org/10.22080/iselt.2023.25121.1049

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass& C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Tseng, C. C. (2019). Senior high school teachers' beliefs about EFL writing instruction.*Taiwan Journal of TESOL, 16*(1), 1-39. https://doi.org/10.30397/TJTESOL.201904_16(1).0001

Wahyuda, M.; Putera, A., &Khuseini, M. (2022). The effectiveness of ProWritingAid toward students' writing skills in senior high school.*Journal of English Teaching and Learning, 1*(1), 25-33.https://doi.org/10.18860/jetl.v1i1.1624

Wang, P. (2013). Can automated writing evaluation programs help students improve their English writing? *International Journal of Applied Linguistics and English Literature, 2*(1), 6–12. https://doi.org/10.7575/ijalel.v.2n.1p.6

Wondim, B. M., Bishaw, K. S., &Zeleke, Y. T. (2023). Effects of teachers' written corrective feedback on the writing achievement of first-year Ethiopian university students. *Education Research International, 2023*(1).https:// doi.org/ 10.1155/ 2023/ 7129978

Wondim, B. M., Bishaw, K. S., &Zeleke, Y. T. (2024). Effectiveness of teachers' direct and indirect written corrective feedback provision strategies on enhancing students' writing achievement: Ethiopian university entrants in focus. *Heliyon, 10*(2).https://doi.org/ 10.1016/j.heliyon.2024.e24279

Woodworth, J., &Barkaoui, K. (2020). Perspectives on using automated writing evaluation systems to provide written corrective feedback in the ESL classroom. *TESL Canada journal, 37*(2), 234-247.https://doi.org/10.18806/tesl.v37i2.1340

Zhang, V. (2017). Student engagement with computer-generated feedback: a case study. *ELT Journal, 71*(3), 317–328. https://doi.org/10.1093/elt/ccw089

Appendix I. Paragraph writing activities for the experimental and control groups

No	Activities	Writing Prompt
1	Activity 1.1	Write a paragraph of about 100 words explaining some of the challenges you encounter when listening to a lecture.
2	Activity 1.2	Read an article about health and fitness on page 28 of your module, and write a short paragraph explaining key insights you gained from the article you have read.
3	Activity 1.3	Reflect on your reading habits over the years and write a paragraph of 150-200 words.
4	Activity 1.4	Write a paragraph of 120 words describing a couple of cultural values you are proud of and explain their importance to the society.
5	Activity 1.5	Write a short paragraph (5-8 lines) explaining how the article "Africa's Wild Animals" on page 53 can attract tourists to Africa (or you can take the opposite view).
6	Activity 1.6	Write a short paragraph interpreting the data on page 70 of your module.