Automated feedback and writing achievement in a foreign language context

Gülnihan Altınay[*]

Selami Aydın[**]

Introduction

Writing is considered one of the most challenging and significant language skills in the foreign language learning and teaching process for several reasons. First, writing skills contribute to developing all other language skills such as reading, listening, speaking, vocabulary, and grammar. With this undeniably significant skill, learners can improve learned structures, increase their knowledge of vocabulary, and notice mistakes and deficiencies in the target language (Çakır, 2010), leading to the overall development of learners' language proficiency and facilitating the way from competence to performance (Hyland, 2022). Second, writing skills significantly affect learners' academic success in general (Harmer, 2004). Furthermore, in terms of social interaction, it is of great importance for learners to convey their feelings and thoughts appropriately and effectively to build a successful communication and cooperation environment.

Feedback has numerous positive effects on learners' writing achievement. First, meaningful development can be seen in learners' grammar, spelling, punctuation, and vocabulary knowledge as they are required to edit and rewrite their drafts multiple times to create an accurate piece of writing (Zaman & Kalam, 2012), leading to more opportunities to practice (Kamberi, 2013). Second, advances in content knowledge can also be acquired through feedback since it shows learners how to arrange ideas and information logically and fluently in written work (Hyland & Hyland, 2006).

Despite its numerous benefits and impacts on the learning process, feedback and its process may cause various problems. First, learners may not always receive timely and regular feedback (Brookhart, 2017). The inconsistent and delayed feedback causes problems in learners' overall writing development due to irregular and scarce guidance (Clariana, 1999). It also creates a problem if the feedback given conflicts with the previous feedback or with the feedback that is provided by a different source (Miranty & Widiati, 2021), leading to an

[*] Research assistant, Istanbul Kültür University, Education Faculty, the Department of English Language Teaching, email: g.altinay@iku.edu.tr, ORCID: https://orcid.org/0000-0002-9568-864X.

[**] Prof. Dr., Istanbul Medeniyet University, Faculty of Education Sciences, the Department of English Language Teaching, email: selami.aydin@medeniyet.edu.tr, ORCID: https://orcid.org/0000-0003-1614-874X

ambiguity which may result in learners being more anxious, less motivated, and confused which hinders their improvement. Moreover, vague and ambiguous feedback may cause higher anxiety and demotivation by misguiding and confusing learners about their errors (Hassan & Mahfoodh, 2017) and creating an uncertain picture of their overall progress. Finally, although it occurs between the teacher and learners, learners generally do not have the opportunity to interact with the teacher about the feedback they receive despite its two-way and dynamic nature (Nassaji, 2016), lowering the potential and expected efficacy of feedback by diminishing the opportunity of revision on written work.

The effects of these problems on writing achievement during the feedback process are of great importance. However, the use of automated feedback may diminish these issues and eliminate the potential undesired consequences of traditional teacher feedback. Regarding the issue of when and how feedback is given, automated feedback offers learners clear, immediate, and individualized feedback according to their needs, regardless of time and place constraints (Jiang & Yu, 2022). With its implementation, students can evaluate their written products whenever and wherever they want to see their mistakes and correct them without being in a traditional classroom environment. Although the use of automated feedback may provide promising solutions to the problems EFL learners experience and benefits to their development of writing achievement, research on the related subject is fairly limited and needs further research, as clarified below.

Previous Research

Research indicates that providing learners with feedback enhances their writing skills, resulting in higher levels of writing achievement. For instance, Fathman and Whalley (1990) conducted a study including 72 intermediate ESL learners in a college composition course to investigate the effects of grammar and content feedback on learners' writing achievement. The results of the study showed that the learners provided with feedback surpassed the learners who received no feedback in terms of grammar and content of their writings. Another study conducted by Ferris and Roberts (2001) to measure the impact of feedback on the writing skills of 72 college-level students showed that the written products of the no-feedback group were less accurate and successful in terms of grammar and sentence structure. On the contrary, the groups who received feedback performed better than their counterparts. Chandler (2003) examined the writings of thirty-one learners in terms of grammatical and lexical errors. While the experimental group received feedback on both content and grammatical errors, the control group did not receive feedback and was only allowed to edit their writing without any specification of the errors. The results of the study showed that the experimental group outperformed the control group at the end of the 14-week study, indicating the positive impacts of feedback on learners' accuracy in writing. To examine the efficacy of written corrective feedback, Elfiyanto and Fukazawa (2021) conducted research including 162 learners from two different high schools. In this three-week-long study, learners are divided into three groups: a teacher-feedback group, a peer-feedback group, and a self-correction group. Learners in each group composed five argumentative essays in 40 minutes and learners' pre-test and post-test essays were used as the tools to obtain results. At the end of the study, the participants of all three groups' accuracy and writing achievement increased.

The results of a fairly limited number of studies show that the implementation of automated feedback in writing evaluation and feedback processes can lead to positive outcomes in learners' writing achievement. Kellogg et al. (2007) investigated the effectiveness of automated feedback by using Criterion to examine the impacts of it on fifty-nine EFL learners' writing achievement. Three groups including a control group and two experimental groups were formed according to the feedback conditions. While one of the experimental groups received continuous automated feedback, the other experimental group received automated feedback intermittently. On the other hand, the control group did not receive feedback completely. The results suggested that while the experimental groups outperformed the control group, the group that received continuous feedback from Criterion was more accurate than the intermittent feedback group. Also, Cheng (2017) conducted a study involving fifty-four EFL learners. Dividing learners into two groups, the experimental group received automated feedback on their reflective essays. According to the results, the control group underperformed the experimental group in terms of writing achievement and overall reflective essay points. Unlike previous studies, Altuntaş (2021) compared the efficacy of automated feedback and teacher feedback on learners' writing achievement. Including ninety-one EFL learners, three groups were formed. While one of the two experimental groups received only automated feedback, the other experimental group received both automated and teacher feedback. The control group received only teacher feedback. The results showed that automated feedback increased learners' overall writing achievement.

Overview of the Current Study

As outlined before, learners' writing achievement is highly influenced by certain factors. However, the feedback process is crucial in learners' writing achievement development. While the number of studies on learners' writing achievement has increased due to the importance they hold, there is a gap in the literature regarding the use of automated feedback and its impacts on learners' writing achievement. Thus, this study aims to provide an understanding of the effects of using automated feedback and discover whether automated feedback and traditional teacher feedback possess different impacts on the learners' writing achievement and seeks answers to the following research questions:

1. Does teacher feedback affect the level of writing achievement among EFL learners?
2. Does automated feedback affect the level of writing achievement among EFL learners?
3. Is there a difference in EFL learners' writing achievement in the use of teacher and automated feedback?

Method

Research Design

This study examines the effects of automated feedback on foreign language learners' writing achievement. First, it analyzes the effects of traditional teacher feedback and automated feedback on language learners' writing achievement. In addition, it examines whether there is a difference in learners' writing achievement. The study adopts a quasi-experimental research design in which already existing groups are used instead of random participant assignment (Creswell, 2014). Although using preexisting groups is considered a limitation due to the lack

of randomization, it is practical and natural (Shadish et al., 2002). As in the experimental research design, the study includes an experimental group that receives treatment and a control group that receives no treatment, helping to examine the effects of the treatment.

Participants

The study included 26 junior students, studying at the ELT department of a foundation university in Turkey. While 18 of the participants were females (69.30%), eight of them were males (30.70%). The mean age of the participants was 20.53 while the minimum age was 18 and the maximum age was 31. All students who participated in the study completed the preparatory school which is mandatory for all students and includes extensive courses that focus on speaking, listening, reading, and writing. Therefore, their English proficiency was intermediate (B1) according to the Common European Framework of Reference (CEFR) standards.

Tools

In the study, two tools were used including demographic information to obtain data on learners' background information such as age, gender, grade, department, current proficiency levels, and weekly hours of English courses. Additionally, two writing tasks prepared by Cambridge ESOL were used to evaluate participants' level of writing achievement.

Procedure

After obtaining the necessary ethics committee permission documents, participants were given thorough information including the objective, procedure, and significance of the study to be conducted. It was particularly emphasized that the study was voluntary. Participants were informed and guaranteed that their information would be used anonymously for scientific purposes only and that their data would be kept completely confidential. To obtain the participants' permission, they were given a consent paper to be signed at the beginning of the data collection process.

The study took place in the Fall semester of the 2023-24 academic year and lasted for six weeks including the 4-week long treatment process and pre-test and post-test employment. Participants were required to complete a writing task each week during the practice process. A pre-existing classroom including 26 students was used and the participants were randomly divided in half and assigned to the control and experimental groups. Later, the pre-test was employed to obtain pre-treatment data. During the treatment process, participants were asked to complete a writing task each week under the control of the teacher. While the control group used pen and paper to complete the writing tasks and evaluated according to the Writing Assessment Scale developed by Cambridge English Qualifications, the experimental group used Write&Improve, an online automated feedback tool designed and developed by Cambridge University that enabled students to receive immediate, automated, and individualized feedback. At the end of the practice process, all participants took the post-test.

At the beginning of the study, the background information questionnaire consisting of six questions was distributed to the participants to obtain background information such as participants' age, gender, grade, English proficiency levels, department, and weekly hours of English classes. To complete the writing achievement task, participants were given 60 minutes.

To obtain valid data on their current writing achievement levels, participants were not permitted to use dictionaries or consult each other about the meanings of unknown or forgotten words.

During the 4-week long practice process, the participants were asked to complete pre-determined writing tasks taken from Write&Improve by Cambridge. The writing tasks were chosen from the Write&Improve Advanced category and participants were required to write between 200-250 words. Before each task, participants were informed about the specific type of essay types including the distinct characteristics and example essays they were required to examine. In addition, the questions of the participants were answered by the teacher before the actual writing task. The control group of the study used pen and paper to complete the tasks while the experimental group used Write&Improve. Therefore, while the control group received traditional teacher feedback, the experimental group received immediate automated feedback from the automated feedback tool, as presented in Table 1.

Weeks	Writing Tasks	Instruction	Type of Feedback	
			Control Group	Experimental Group
Week 1	Writing an opinion essay about the main problems of Turkish teenagers by describing causes and providing solutions	Characteristics of an opinion essay, descriptive essay, and review	Teacher Feedback	Automated Feedback
Week 2	Writing a descriptive essay about a way of relaxing by implying its importance in today's world.	How to express an opinion		
Week 3	Writing a review describing the best and worst films you have seen and giving reasons for these choices.	How to support opinions and provide reasons		
Week 4	Writing an essay discussing two of the areas in the notes given, identifying which area has benefited more from the internet and giving reasons for the specific choice.	Discussions		

Table 1. The practice process

Data Analysis

IBM SPSS Statistics 21.0 was used to analyze the data. First, the maximum, minimum, and mean of the participants' ages were analyzed. Then, the participants' gender percentages and numbers were computed. Later, through Cronbach's Alpha, the reliability of the tools employed in the study was computed as well as Varimax rotation for both the pre-test and post-test. The reliability coefficient value of the writing achievement tasks was calculated as .70 for the pre-test and .73 for the post-test. The percentage of variance of the writing achievement tasks was 52.63% for the pre-test and 55.60% for the post-test.

For further analyses, non-parametric methods were employed since the number of participants was small and could not be regarded as a normal distribution. Descriptive analysis, the Wilcoxon analysis, and Mann-Whitney U analysis were used to analyze the data. First, descriptive analysis was conducted to find the mean scores of the pre-test and post-test results. Then, the Wilcoxon analysis was employed to examine whether there were any changes between the pre-test and post-test results of the control and experimental groups' writing achievement levels after. Later, the Mann-Whitney U analysis was carried out to compare whether there were any significant differences between the scores of the groups.

Results

The Effects of Teacher Feedback on Learners' Writing Achievement

As shown in Table 2, while participants' overall writing achievement scores had a mean of 13.31 in the pre-test, it decreased to 12.54 in the post-test. Furthermore, it presents the control group's mean, minimum, and maximum scores and standard deviation values of content achievement, communicative achievement, organization, and language use. As seen in the table, the control group's content achievement mean score was 3.31 before the implementation and 2.92 after the implementation, indicating that the control group's content achievement decreased after the implementation process. Concerning communicative achievement scores, the participants in the control group had a mean of 3.15 in the pre-test scores and 3.15 in the post-test, indicating no difference in participants' mean scores of communicative achievement between the pre-test and post-test implementation. For organization, the participants in the control group had a mean of 3.77 in the pre-test and 3.62 in the post-test. Finally, the participants' language use scores had a mean of 3.08 in the pre-test and 2.85 in the post-test, showing a decrease in the participants' language use scores after the implementation.

Content	Tests	Mean	Std. Deviation	Minimum	Maximum
Overall Achievement	Pre-test	13.31	2.72	8	18
	Post-test	12.54	1.94	9	16
Content Achievement	Pre-test	3.31	.75	2	5
	Post-test	2.92	.76	2	4
Communicative Achievement	Pre-test	3.15	.80	2	5
	Post-test	3.15	.69	2	4
Organization	Pre-test	3.77	1.17	2	5
	Post-test	3.62	.65	3	5
Language Use	Pre-test	3.08	.87	2	5
	Post-test	2.85	.69	2	4

Table 2. The mean scores for the control group (Wilcoxon test, n=13)

Table 3 compares the pre-test and post-test scores for the overall writing achievement of the control group including the number of positive and negative ranks, mean ranks, sum of ranks, and significance level of the comparison. According to the results, five of the participants' writing achievement scores decreased with a mean rank of 6.40 while four of the participants' scores increased with a mean rank of 3.25. In addition, four participants' writing

achievement scores showed no difference after the implementation. With a significance level of .26, it can be stated that there is no significant difference in participants' writing achievement scores between the pre-test and post-test. Furthermore, it compares the pre-test and post-test scores for content achievement, communicative achievement, organization, and language use scores of the control group including the number of positive and negative ranks, mean ranks, sum of ranks, and significance level of the comparison. As seen in the table, four of the participants' content achievement scores decreased with a mean rank of 3.25 while one participant's content achievement score increased with a mean rank of 2.00. In addition, it is seen that eight of the participants' content achievement scores did not show any difference. Since the significance value was .13, there is no significant difference in the participants' content achievement scores between the pre-test and post-test. In addition, the communicative achievement scores of three participants were lower in the post-test with a mean rank of 10.50 while three participants' communicative achievement scores increased with a mean rank of 10.50. In addition, seven of the control group participants' communicative achievement scores remained the same as in the pre-test. With a significance level of 1.0, it can be stated that no significant difference was encountered between the pre and post-test scores. In terms of organization, five participants' scores decreased with a mean rank of 6.50 while the organization scores of five participants increased with a mean rank of 4.50 after the implementation. Also, the scores of the three participants showed no difference in the post-test. Although little differences can be seen according to the results, there is no significant difference between the pre-test and post-test results. Lastly, six participants' language use scores decreased with a mean rank of 5.00 while the language use scores of three participants increased with a mean rank of 5.00 after the practice process. However, no significant difference was encountered between the pre and post-test scores due to the significance level of .32.

Criteria		N	Mean Ranks	Sum of Ranks	Asymp. Sig. (2-tailed)
Overall Achievement	Negative Ranks	6	5.00	30.00	.26
	Positive Ranks	4	3.25	13.00	
	Ties	4			
Content Achievement	Negative Ranks	4	3.25	13.00	.13
	Positive Ranks	1	2.00	2.00	
	Ties	8			
Communicative Achievement	Negative Ranks	3	3.50	10.50	1.0
	Positive Ranks	3	3.50	10.50	
	Ties	7			
	Negative Ranks	5	6.50	32.50	.59

Organization	Positive Ranks	5	4.50	22.50	
	Ties	3			
	Negative Ranks	6	5.00	30.00	
Language Use	Positive Ranks	3	5.00	15.00	.32
	Ties	4			

Table 3. The comparison of pre and post-test scores for the control group (Wilcoxon test, n=13)

The Effects of Automated Feedback on Writing Achievement

Table 4 presents the overall writing achievement scores of the participants in the experimental group including the mean score, standard deviation as well as minimum and maximum means. According to the table, while participants' overall writing achievement scores had a mean of 13.54 in the pre-test, it was calculated as 14.15 in the post-test, indicating an increase after the implementation. Furthermore, it shows the content achievement, communicative achievement, organization and language use scores of the experimental group including the mean score, and standard deviation as well as minimum and maximum means. According to the table, while the mean score of the participants' content achievement was calculated as 3.62 in the pre-test, it was 3.46 in the post-test, displaying a decrease after the implementation. For communicative achievement, while the mean score of the participants was 3.15 in the pre-test, it was calculated as 3.23 in the post-test. In terms of organization, while the mean score of the participants was 3.92 in the pre-test, it increased to 4.08 after the implementation. Finally, the participants' mean score of language use was calculated as 3.92 in the pre-test, it increased to 4.08 after the implementation.

Criteria	Tests	Mean	Std. Deviation	Minimum	Maximum
Overall Achievement	Pre-test	13.54	2.14	11	18
	Post-test	14.15	2.79	10	18
Content Achievement	Pre-test	3.62	.87	2	5
	Post-test	3.46	.97	2	5
Communicative Achievement	Pre-test	3.15	.69	2	5
	Post-test	3.23	.92	2	5
Organization	Pre-test	3.92	.64	3	5
	Post-test	4.08	.86	3	5
Language Use	Pre-test	3.00	.71	2	4
	Post-test	3.38	.96	2	5

Table 4. The mean scores for the experimental group (Wilcoxon Test, n=13)

Table 5 compares the pre-test and post-test scores for the overall writing achievement of the experimental group presenting the number of positive and negative ranks, mean ranks, sum of ranks, and significance level of the comparison. According to the results, five of the

participants' writing achievement scores decreased with a mean rank of 4.40 while six of the participants' scores increased with a mean rank of 7.33. Also, the two participants' overall writing achievement scores did not change after the implementation. Although there is a slight increase in participants' overall writing achievement scores, there is no significant difference between the pre-test and post-test with a significance level of .32. Furthermore, it presents the comparison of the pre-test and post-test scores for the participants' content achievement, communicative achievement, organization, and language use scores presenting the number of positive and negative ranks, mean ranks, sum of ranks, and significance level of the comparison. As can be seen in the table, while five participants' content achievement scores decreased after the implementation, four participants' content achievement scores increased with a mean rank of 5.00. Moreover, four participants' scores did not show any difference in the post-test. The significance level was calculated as .75. In terms of communicative achievement, five of the participants' communicative achievement scores decreased with a mean rank of 4.00 while four of the participants' scores increased with a mean rank of 6.25. Also, four participants' communicative achievement scores remained the same after the implementation. Although there is an increase in participants' communicative achievement scores, there is no significant difference between the pre-test and post-test with a significance level of .75. In the matter of organization, two participants' organization scores decreased with a mean rank of 3.50 while four participants' scores increased with a mean rank of 3.25. Also, the organization scores of the seven participants remained the same in the post-test. Although there is an increase in participants' organization scores, there is no significant difference between the pre-test and post-test with a significance level of .41. Finally, two participants' language use scores decreased with a mean rank of 5.00 while seven participants' scores increased with a mean rank of 5.00. Furthermore, the language use scores of the four participants did not show any change in the post-test. Although there is an increase in participants' language use scores, there is no significant difference between the pre-test and post-test scores with a significance level of .10.

Criteria	Ranks	N	Mean Ranks	Sum of Ranks	Asymp. Sig. (2-tailed)
Overall Achievement	Negative Ranks	6	7.33	44.00	.75
	Positive Ranks	2	7.33	44.00	
	Ties	6			
Content Achievement	Negative Ranks	5	5.00	25.00	.75
	Positive Ranks	4	5.00	20.00	
	Ties	4			
Communicative Achievement	Negative Ranks	5	4.00	20.00	.75
	Positive Ranks	4	6.25	25.00	
	Ties	4			
	Negative Ranks	2	3.50	7.00	.41

Organization	Positive Ranks	4	3.50	14.00	
	Ties	7			
	Negative Ranks	2	5.00	10.00	
Language Use	Positive Ranks	7	5.00	35.00	.10
	Ties	4			

Table 5. Comparison of the scores for the experimental group (Wilcoxon test, n=13)

The difference in writing achievement for teacher and automated feedback

Table 6 below presents the pre-test overall writing achievement scores of the control and experimental groups including the mean rank, sum of ranks, and significance level. As can be seen in the table, the participants in the control group had a mean rank of 13.38 while the experimental group was 13.62. As the significance level was calculated as .94, it can be stated that there is no significant difference between the groups in the pre-test.

	Group	Mean Rank	Sum of Ranks	Asymp. Sig. (2-tailed)
Pre-test	Control Group	13.38	174.00	.94
	Experimental Group	13.62	177.00	

Table 6. The comparison of the groups (Mann-Whitney U Test)

Table 7 provides the pre-test scores for content achievement, communicative achievement, organization, and language use scores of the control and experimental groups indicating the mean rank, sum of ranks, and significance level. According to the table, in terms of content achievement, while the control group had a mean rank of 12.08, the experimental group's mean rank was 14.92. However, there is no significant difference between the groups as the significance level was calculated as .30. For communicative achievement, both the control and experimental groups' mean rank was 13.50 with a significance level of 1.00, indicating that there is no significant difference between the control and experimental groups. In addition, the mean rank of the control group's organization scores was calculated as 13.15 while the experimental group received a mean rank of 13.85 in the pre-test. The significance level was .81; therefore, there is no significant difference between the groups. Lastly, the control group's language use scores had a mean rank of 13.62 while the experimental group had a mean rank of 13.38 with a significance level of .93.

	Category	Group	Mean Rank	Sum of Ranks	Asymp. Sig. (2-tailed)
Pre-test	Content Achievement	Control Group	12.08	157.00	0.30
		Experimental Group	14.92	194.00	
	Communicative Achievement	Control Group	13.50	175.50	1.00
		Experimental Group	13.50	175.50	
	Organization	Control Group	13.15	171.00	.81

	Experimental Group	13.85	180.00	
	Control Group	13.62	177.00	
Language Use	Experimental Group	13.38	174.00	.93

Table 7. The comparison of pre-test scores for the groups (Mann-Whitney U Test)

Table 8 below presents the comparison of overall writing achievement of the control and experimental groups including the mean rank, sum of ranks, and significance level values. According to the table, the control group's mean rank was 11.19 while the experimental group's was calculated as 15.81. However, the difference cannot be considered significant.

	Group	Mean Rank	Sum of Ranks	Asymp. Sig. (2-tailed)
Post-test	Control Group	11.19	145.50	.12
	Experimental Group	15.81	205.50	

Table 8. Post-test overall writing achievement scores of the groups (Mann-Whitney U Test)

Table 9 provides the post-test scores for content achievement, communicative achievement, organization, and language use scores of the control and experimental groups indicating the mean rank, sum of ranks, and significance level. According to the table, in terms of content achievement, while the control group had a mean rank of 11.27, the experimental group's mean rank was 15.73. Although there is an increase in the experimental group's scores, there is no significant difference between the groups as the significance level was calculated as .12. For communicative achievement, the control group's mean rank was 13.27 while it was 13.73 for the experimental group with a significance level of .87, indicating that there is no significant difference between the control and experimental groups. In addition, the mean rank of the control group's organization scores was calculated as 11.50 while the experimental group's was 15.50 in the post-test. The significance level was .15, indicating no significant difference between the groups. Lastly, the control group's language use scores had a mean rank of 11.27 while the experimental group had a mean rank of 15.73 with a significance level of .13. Although there is an increase in the experimental group's language use in the post-test, the difference cannot be considered meaningful.

	Category	Group	Mean Rank	Sum of Ranks	Asymp. Sig. (2-tailed)
Post-test	Content Achievement	Control Group	11.27	146.50	.12
		Experimental Group	15.73	204.50	
	Communicative Achievement	Control Group	13.27	172.50	.87
		Experimental Group	13.73	178.50	
	Organization	Control Group	11.50	149.50	.15
		Experimental Group	15.50	201.50	

Language Use	Control Group	11.27	146.50	
	Experimental Group	15.73	204.50	.13

Table 9. The Comparison of post-test scores for the groups (Mann-Whitney U Test)

Conclusion

The study concluded that teacher feedback has no facilitative effects on EFL writers' overall writing achievement including aspects such as content achievement, communicative achievement, organization, and language use. In other words, EFL writers become less accurate and fluent after getting traditional teacher feedback. However, automated feedback enhances EFL writers' writing performance. The use of automated feedback increases EFL learners' writing achievement, including content achievement, communicative achievement, organization, and language use. Therefore, they become more proficient, accurate, and fluent writers. In conclusion, when the use of teacher feedback and automated feedback is compared in terms of their effects on EFL writers' writing achievement, the use of automated feedback is more effective than traditional teacher feedback, helping them to become more proficient writers in terms of both accuracy and fluency.

The results of a focus group study indicate that the abovementioned conclusions may relate to several factors. In terms of writing achievement, participants who received teacher feedback state that the teacher needs to notice their success or failure, so they write more cautiously. However, they experience problems with the organization, stating that since their time is limited, making corrections in the text on a sentence or paragraph basis is time-consuming and difficult, making it complicated to rearrange the already written text.

Several pedagogical implications can be made considering the present study and previous literature. First, teacher feedback does not contribute to learners' writing achievement. This result does not match the majority of earlier literature. According to the literature, teacher feedback contributes to learners' writing performance, improving both the accuracy and fluency of the written work by indicating strengths and weaknesses and guiding the way to the improvement of overall writing achievement (Chandler, 2003; Elfiyanto & Fukazawa, 2021; Fathman & Whalley, 1990; Ferris & Roberts, 2001). There may be several reasons for this conflict. First, the research indicates that most students prefer explicit feedback over exploratory feedback, believing that explicit feedback received from the teacher is a more understandable and faster way to correct their errors (Bulut, 2019). Another reason may be the limitations of time and place. In other words, learners receive feedback at a time and place determined by the teacher, making them dependent on the teacher and inhibiting them from receiving instant feedback (Waer, 2023). Lastly, learners do not have the chance to edit their written work according to the corrections and suggestions given by the teacher (Lee, 2007). According to the current study, the use of automated feedback increases EFL learners' writing performance in terms of accuracy and fluency, by providing immediate and individualized feedback during the writing process. This result is compatible with the limited number of previous studies in the literature (Altuntaş, 2021; Benali, 2021; Cheng, 2017; Kellogg & Raulerson, 2007). Automated feedback helps learners improve their writing skills in several ways. First, it provides learners with immediate and continuous feedback (Benali, 2021).

Second, by giving them chances for multiple revisions, learners can fix their errors right after receiving feedback, making the learning process faster and easier (Benali, 2021). Third, since learners do not interact with an authority figure, they can write freely without fear of judgment (Benali, 2021). Considering the significant effects of psychological factors on learners' writing performance, automated feedback allows learners to write in a safe environment. As a result, automated feedback tools can play an important role in improving learners' writing skills and provide learners with a confident writing environment by meeting their individual learning needs. Therefore, such technological supports can offer an effective way to increase learners' writing performance and optimize their learning processes in foreign language writing.

Several pedagogical implications can be made according to the findings of the present study. Considering the significant development in learners' writing achievement, the use of automated feedback can facilitate the writing process. In addition, the benefits of automated feedback tools for EFL teachers in terms of time management and workload should not be ignored. Therefore, teachers should utilize automated feedback tools to improve students' writing skills including accuracy and fluency, and provide a stress-free learning environment.

The present study has several limitations. First, this study is limited to a total of 26 first-year university students studying in the ELT program at a foundation university, including 13 students in the control group and 13 students in the experimental group. Therefore, the results obtained from this study cannot be generalized in terms of participants, time, place, and process. Second, the implementation period is limited to four weeks. A longer implementation period may provide different findings. Therefore, examination of any significant differences that may exist at baseline between control and experimental groups is limited. Third, this study adopts a quasi-experimental study design due to the use of pre-existing groups instead of random assignment of participants. Finally, only Write&Improve was used as an automated feedback tool in the present study.

According to the results of this study, several recommendations for further research can be made. First, more research should be conducted on the effects of automated feedback on EFL learners' writing achievement. The sample diversity of the study should be increased by including more students from different universities and with different cultural backgrounds. More detailed studies can be conducted on the ease of use and effectiveness of tools such as Write&Improve. Furthermore, to further examine the quality and impact of feedback, a content analysis should be conducted to compare the impacts of both teacher and automated feedback on learners' writing.

Ethics Committee Approval

The authors confirm that the ethical approval was obtained from the Ethical Committee of Educational Sciences (Approval Number: 2023/159 Date: 07.12.2023)

Acknowledgment

This article is a version of the first author's M.A. thesis advised by the second author.

References

Altuntaş, T. (2021). *The use of automated writing evaluation in EFL writing classes: Effects on achievement, motivation and self-efficacy* [Master's Thesis, Gazi University].

Benali, A. (2021). The impact of using automated writing feedback in ESL/EFL classroom contexts. *English Language Teaching*, *14*(12), 189. https://doi.org/10.5539/elt.v14n12p189

Brookhart, S. M. (2017). *How to give feedback to your students* (2nd ed.). ASCD. https://books.google.com/books?id=CYzgEAAAQBAJ

Bulut, S. E. (2019). *The effects of automated writing evaluation on EFL students' writing achievement and motivation towards writing* [Master's Thesis, Istanbul University]. http://acikerisim.iuc.edu.tr/xmlui/handle/20.500.12831/4875

Çakır, İ. (2010). Yazma becerisinin kazanılması yabancı dil öğretiminde neden zordur? [Why is it difficult to acquire writing skills in foreign language teaching?]. *Erciyes University Journal of Social Sciences Institute*, *1*(28), 1–12. https://dergipark.org.tr/en/pub/erusosbilder/issue/23762/253280

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education*, *34*, 18–27. https://doi.org/10.1016/j.iheduc.2017.04.002

Clariana, R. B. (1999, February). Differential memory effects for immediate and delayed feedback: A delta rule explanation of feedback timing effects. *Association for Educational Communications and Technology Annual Convention*.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Elfiyanto, S., & Fukazawa, S. (2021). Three written corrective feedback sources in improving Indonesian and Japanese students' writing achievement. *International Journal of Instruction*, *14*(3), 433–450. https://doi.org/10.29333/iji.2021.14325a

Fathman, A. K., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178–190). Cambridge University Press.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes How explicit does it need to be? *Journal of Second Language Writing*, *10*(3), 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Harmer, J. (2004). *How to teach writing*. Pearson/Longman.

Hassan, O., & Mahfoodh, A. (2017). "I feel disappointed": EFL university students' emotional responses towards teacher written feedback. *Assessing Writing*, *31*, 53–72. https://doi.org/10.1016/j.asw.2016.07.001

Hyland, K. (2022). *The Routledge Handbook of Second Language Acquisition and Writing* (R. M. Manchón & C. Polio, Eds.; 2nd ed.). Routledge.

Hyland, K., & Hyland, F. (2006). *Feedback in Second Language Writing*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524742

Jiang, L., & Yu, S. (2022). Appropriating automated feedback in L2 writing: Experiences of Chinese EFL student writers. *Computer Assisted Language Learning*, *35*(7), 1329–1353. https://doi.org/10.1080/09588221.2020.1799824

Kamberi, L. (2013). The significance of teacher feedback in EFL writing for tertiary level foreign language learners. *Procedia - Social and Behavioral Sciences*, *70*, 1686–1690. https://doi.org/10.1016/J.SBSPRO.2013.01.241

Kellogg, R. T., & Raulerson, B. A. (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review*, *14*(2), 237–242. https://doi.org/10.3758/BF03194058

Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, *12*(3), 180–198. https://doi.org/10.1016/j.asw.2008.02.003

Miranty, D., & Widiati, U. (2021). Automated writing evaluation (AWE) in higher education: Indonesian EFL students' perceptions about Grammarly use across student cohorts. *Pegem Egitim ve Ogretim Dergisi*, *11*(4), 126–137. https://doi.org/10.47750/pegegog.11.04.12

Nassaji, H. (2016). Anniversary article interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, *20*(4), 535–562. https://doi.org/10.1177/1362168816644940

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin.

Waer, H. (2023). The effect of integrating automated writing evaluation on EFL writing apprehension and grammatical knowledge. *Innovation in Language Learning and Teaching*, *17*(1), 47–71. https://doi.org/10.1080/17501229.2021.1914062

Zaman, M., & Kalam, A. (2012). Feedback in EFL writing at tertiary level: Teachers' and learners' perceptions. *ASA University Review*, *6*(1).