**PROCEEDINGS**
International Conference on Education
Faculty of Tarbiyah | UIN Syekh Wasil Kediri
E-ISSN: 3025-6828

Volume 4, 2026

# Digital Journaling Through Google Docs to Improve Students' Writing Achievement at an Islamic Boarding School

**Fitri Handayani[1], Wildania Maghfurin[1], Arina Chusnatayaini[2]**
[1]Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung;
[2]Universitas STRADA Indonesia
Correspondence e-mail: fitriakunsekolah@gmail.com

**Abstract:**
Traditional Islamic boarding schools are often associated with conventional learning practices, while digital learning activities are increasingly used to support students' academic development. This study examined the potential value of digital journaling through Google Docs in supporting university students' English writing achievement at Subulussalam Islamic Boarding School. A quantitative approach with a quasi-experimental non-equivalent control group design was employed involving 30 university students, with 15 students in the experimental group and 15 in the control group. Data were collected through pre-test and post-test writing assessments and scored using an analytical writing rubric. The data were analyzed using descriptive statistics, Shapiro–Wilk and Levene's tests, and the Mann–Whitney U test. The experimental group's mean score increased from 50.13 to 81.47, while the control group's mean score increased from 36.27 to 60.20. The experimental group therefore showed a greater descriptive improvement (31.34 points) than the control group (23.93 points). The Mann–Whitney U test indicated significant differences between the groups at both pre-test ($p = 0.005$) and post-test ($p < 0.001$), showing that the groups were not equivalent at baseline. Therefore, although digital journaling was associated with higher post-test writing achievement and greater descriptive improvement, the findings do not provide conclusive evidence that the intervention alone caused the observed difference. The study highlights the potential of Google Docs-supported journaling as a continuous writing, feedback, and revision activity in an Islamic boarding school context.

**Keywords: Digital Journaling, Google Docs, Writing achievement, English Language Learning, Islamic Boarding School**

## Introduction

Traditional Islamic boarding schools for university students are commonly associated with religious education and conventional teaching practices. However, Subulussalam Islamic Boarding School offers a different learning experience by integrating English language learning with digital journaling activities through Google Docs (Nur & Ramadhani, 2025; Zahrok & Rahmawati, 2025). The integration of technology reflects the institution's effort to maintain its Islamic educational values while responding to the needs of twenty-first-century learning. As an Islamic boarding school that accommodates university students from different higher education institutions in Tulungagung, Subulussalam also supports students' academic development, including their English language proficiency. In this context, writing is an important skill because university students need to express ideas, complete academic tasks, and communicate information effectively in written English (Hyland, 2022).

Despite its importance, writing in English remains challenging for many EFL learners. Students need to generate and develop ideas, organize paragraphs coherently, use appropriate grammatical structures, select suitable vocabulary, and express their ideas clearly. Preliminary observations at Subulussalam Islamic Boarding School showed that many students experienced difficulties in generating ideas, organizing coherent paragraphs, applying appropriate grammatical structures, and expressing their thoughts confidently in written English. They also had limited opportunities to practice writing continuously outside classroom activities. These conditions may result in low writing engagement and limited improvement in writing achievement. Therefore, students need a learning activity that provides regular writing practice, opportunities for feedback, and opportunities to revise their work. Digital journaling through Google Docs offers a relevant alternative because students can write regularly, receive feedback, revise their writing, and monitor their progress through a familiar digital platform. It can also support learner autonomy and digital literacy, which are increasingly important in higher education (Iftanti, Madayani, & Ningrum, 2024).

Previous studies have reported positive effects of journal writing and digital technology on students' writing development. Sa'diyah and Nabhan (2021) reported that Google Docs facilitated collaborative writing and continuous revision through online interaction. More recent studies have also indicated that Google Docs can support collaborative writing, learner engagement, and writing quality in EFL classrooms (Sembiring et al., 2025). Other studies have shown that journaling activities can help students generate ideas, organize paragraphs, improve grammatical accuracy, develop vocabulary, and build confidence in writing (Alwahoub, Jomaa, & Azmi, 2022). In addition, technology-supported language learning can contribute to students' digital literacy and provide more flexible learning opportunities in higher education (Iftanti et al., 2024). These findings indicate that journaling and digital technology have considerable potential to support English writing development (Valizadeh, 2022).

However, most previous studies have examined journal writing or digital writing activities in formal educational settings, such as schools and universities (Kurniawati, 2022). Limited attention has been paid to the use of digital journaling in Google Docs in traditional Islamic boarding schools among university students. This context is important because students in Islamic boarding schools have a particular learning environment in which academic, religious, and technological experiences are integrated. Although previous research has established the potential benefits of journaling and Google Docs, empirical evidence concerning their implementation in this specific educational context remains limited. Therefore, further investigation is needed to determine whether digital

journaling can effectively support the writing development of university students in an Islamic boarding school.

Based on this research gap, this study investigates the effectiveness of digital journaling activities through Google Docs in improving the writing achievement of university students at Subulussalam Islamic Boarding School. The novelty of this study lies in the implementation of digital journaling within a traditional Islamic boarding school context involving university students. Unlike studies that mainly focus on conventional journaling or digital writing in formal educational institutions, this study combines continuous digital writing practice, lecturer feedback, revision, and reflective learning through Google Docs while maintaining the educational values of the Islamic boarding school. The study is expected to contribute theoretically by extending the discussion of technology-enhanced writing instruction to a non-formal Islamic educational context and practically by providing an alternative strategy for English teachers, lecturers, and Islamic boarding school educators to support students' writing achievement, digital literacy, learner autonomy, and engagement in English learning.

## Methods

This study employed a quantitative approach using a quasi-experimental design with a non-equivalent control group. The research was conducted at Subulussalam Islamic Boarding School, Tulungagung, East Java, Indonesia. The participants were 30 university students from the Marhalah Ula and Marhalah Tsani classes who participated in additional English learning activities. They were selected using purposive sampling based on the criteria established by the researcher. The participants consisted of 15 students in the experimental group and 15 students in the control group. The experimental group received digital journaling activities through Google Docs, while the control group received conventional writing instruction. The data source was the writing performance of the participating students, and the data consisted of quantitative writing scores obtained from the pre-test and post-test (Creswell & Creswell, 2023).

The main research instrument was a writing test administered as a pre-test and post-test. In both tests, students were required to write an English paragraph based on a given topic. Their writing was assessed using an analytic scoring rubric covering content, organization, vocabulary, language use, and mechanics, informed by recent research on analytic EFL writing assessment (Phuong, Phan, & Le, 2023). Two independent raters assessed the students' writing, and the inter-rater correlation coefficient was 0.87, indicating high reliability. The study was conducted over six meetings. In the first meeting, both groups completed the pre-test, and the experimental group was introduced to digital journaling through Google Docs. During the second to fifth meetings, students in the experimental group completed digital journal entries based on topics related to their daily experiences and academic life. They received lecturer feedback through the comment feature, revised their writing, and resubmitted the revised versions. The control group received conventional writing instruction without Google Docs or digital journaling. In the sixth meeting, both groups completed the post-test.

The collected writing scores were analyzed using SPSS. Descriptive statistics were used to summarize the pre-test and post-test scores. The Shapiro–Wilk test was used to examine data normality, while Levene's test was used to examine homogeneity of variance between the groups. The Shapiro–Wilk results showed that all significance values were above 0.05, indicating normally distributed data. Levene's test showed a significance value of 0.061 for the pre-test and 0.011 for the post-test; therefore, the post-test data did not meet the homogeneity assumption. The significance level was set at 0.05.

## Results and Discussion

### Findings

This study examined the writing achievement of university students who participated in digital journaling activities through Google Docs and compared their performance with students who received conventional writing instruction. The findings are presented in five parts: descriptive results of students' writing scores, comparison of score improvement, normality test, homogeneity test, and hypothesis testing. The data were obtained from the pre-test and post-test writing scores of 30 students, consisting of 15 students in the experimental group and 15 students in the control group. The experimental group participated in digital journaling activities through Google Docs, while the control group received conventional writing instruction.

1. Descriptive Results of Students' Writing Achievement

Students' writing achievement was measured through a pre-test and a post-test. The pre-test was administered before the treatment to identify the students' initial writing performance, while the post-test was administered after the treatment to measure their writing achievement following the learning process. The same writing assessment procedure was applied to both the experimental and control groups. The results provide information about the students' writing performance before and after the learning process.

2. Experimental Group

The experimental group's writing achievement increased from the pre-test to the post-test. The total pre-test score was 752, resulting in a mean score of 50.13. The individual pre-test scores ranged from 23 to 70. After the treatment involving digital journaling activities through Google Docs, the total post-test score increased to 1,222, with a mean score of 81.47. The post-test scores ranged from 70 to 93. Thus, the difference between the pre-test and post-test mean scores was 31.34 points.

The individual scores also show that improvement occurred across the experimental group. The student who obtained the lowest pre-test score, 23, obtained 70 in the post-test, representing an increase of 47 points. Another student increased from 40 to 93, representing an increase of 53 points. A student who obtained 30 in the pre-test achieved 87 in the post-test, representing an increase of 57 points. A student who initially obtained 53 increased to 93, while another student increased from 60 to 80. These individual score changes demonstrate that the increase in writing performance was observed among students with different levels of initial achievement.

The experimental group also showed improvement among students who had relatively higher initial scores. For example, a student whose pre-test score was 70 obtained 90 in the post-test, while a student who initially obtained 67 obtained 80 and another student with the same initial score obtained 90. Students with pre-test scores of 47 increased to 77 and 73, respectively. A student with a pre-test score of 50 obtained 90, while another student with the same initial score obtained 70. Therefore, the increase was observed across the range of initial scores rather than being limited only to students with the lowest pre-test scores.

The lowest score in the experimental group increased from 23 in the pre-test to 70 in the post-test. The highest score also increased from 70 to 93. The difference in the minimum scores was therefore 47 points, while the difference in the maximum scores was 23 points. These changes indicate that the overall distribution of writing scores moved upward after the treatment. In addition, all 15 students in the experimental group showed an increase from their pre-test to post-test scores.

3. Control Group

The individual scores indicate that improvement also occurred in the control group, although the amount of improvement varied among students. For example, one student increased from 33 to 73, representing a 40-point increase, while another increased from 27 to 73, representing a 46-point increase. A student who initially obtained 37 increased to 73, while another student with a pre-test score of 37 obtained 83 in the post-test. The latter represented the highest increase in the control group, with a difference of 46 points.

Several students in the control group demonstrated smaller increases. One student increased from 43 to 53, while another increased from 43 to 50. A student with a pre-test score of 37 increased to 43, and another increased from 23 to 40. Two students did not show an increase: one remained at 50 in both tests, and another remained at 40. Thus, unlike the experimental group, where all students showed an increase, not all students in the control group demonstrated higher post-test scores.

The lowest score in the control group increased from 23 in the pre-test to 40 in the post-test, representing a 17-point increase. The highest score increased from 50 to 83, representing a 33-point increase. These results show that writing achievement in the control group also improved during the study period, although the pattern of individual improvement was less consistent than in the experimental group.

4. Comparison of the Two Groups

A comparison of the descriptive results shows that both groups experienced an increase in their writing scores. However, the experimental group had a higher mean score than the control group at both measurement points. The experimental group increased from 50.13 in the pre-test to 81.47 in the post-test, while the control group increased from 36.27 to 60.20. The experimental group therefore showed a descriptive mean increase of 31.34 points, whereas the control group showed an increase of 23.93 points. The difference between the two mean improvements was 7.41 points.

Table 1. Summary of Students' Writing Scores

| Group | N | Pre-test Mean | Post–test Mean | Mean Improvement |
|---|---|---|---|---|
| Experimental | 15 | 50.13 | 81.47 | 31.34 |
| Control | 15 | 36.27 | 60.20 | 23.93 |

Source: Authors' data

Table 1 summarizes the pre-test and post-test writing scores of the experimental and control groups. The experimental group obtained a mean pre-test score of 50.13 and a mean post-test score of 81.47, resulting in a mean improvement of 31.34 points. Meanwhile, the control group increased from 36.27 in the pre-test to 60.20 in the post-test, representing a mean improvement of 23.93 points.

The post-test mean difference between the two groups was 21.27 points. The experimental group obtained a post-test mean of 81.47, whereas the control group obtained a mean of 60.20. The minimum post-test score was also higher in the experimental group, with a score of 70 compared with 40 in the control group. The maximum post-test score was 93 in the experimental group and 83 in the control group.

The individual score data further show that all 15 students in the experimental group improved from pre-test to post-test. In the control group, 13 students improved, while two students obtained the same score in both tests. The median score of the experimental group increased from 50 to 80, while the median score of the control group increased from 37 to 60. The median improvement was 26 points in the experimental group and 23 points in the control group. These descriptive patterns show that the

experimental group demonstrated a broader and more consistent upward movement in writing scores.

Overall, the descriptive data indicate that the experimental group achieved higher writing scores following the treatment and experienced a larger average increase than the control group. However, the descriptive comparison should be considered together with the statistical tests because the two groups were not statistically equivalent at the beginning of the study.

5. Normality Test

Before conducting the hypothesis test, the pre-test and post-test scores were examined to determine whether the data met the assumption of normality. The Shapiro–Wilk test was used because the number of participants in each group was fewer than 50. The test was conducted separately for the experimental and control groups at both the pre-test and post-test stages.

The Shapiro–Wilk test was conducted to examine the normality of the writing scores. As presented in Table 2, all significance values were above 0.05, indicating that the pre-test and post-test scores in both groups were normally distributed.

Table 2. Shapiro–Wilk Normality Test

| Data | Sig. | Decision |
|---|---|---|
| Experimental pre-test | 0.664 | Normally distributed |
| Control pre-test | 0.660 | Normally distributed |
| Experimental post-test | 0.078 | Normally distributed |
| Control post-test | 0.140 | Normally distributed |

Source: Authors' data

The results indicate that none of the four data sets showed a statistically significant departure from normality based on the Shapiro–Wilk test. The experimental group's pre-test scores produced the highest significance value of 0.664, followed by the control group pre-test at 0.660. The post-test scores also had significance values above 0.05, with 0.078 for the experimental group and 0.140 for the control group. Therefore, the normality criterion was met for both groups at both measurement points.

The normality results provide information about the distribution of the writing scores before the comparison between the two groups. Although the four sets of scores were normally distributed according to the reported Shapiro–Wilk results, the subsequent analysis also considered the homogeneity of variance.

6. Homogeneity Test

Levene's test was conducted to examine the homogeneity of variance between the experimental and control groups. As shown in Table 3, the pre-test variance was homogeneous ($p = 0.061$), whereas the post-test variance was not homogeneous ($p = 0.011$).

Table 3. Levene's Test of Homogeneity

| Measurement | Sig. | Decision |
|---|---|---|
| Pre-test | 0.061 | Homogeneous |
| Post-test | 0.011 | Not homogeneous |

Source: Authors' data

The pre-test significance value of 0.061 was higher than 0.05. Based on the criterion used in the study, this indicates that the pre-test variances of the experimental and control groups met the homogeneity assumption. In contrast, the post-test significance value of 0.011 was lower than 0.05, indicating that the post-test variances were not homogeneous.

The difference between the pre-test and post-test homogeneity results is important for understanding the statistical procedure reported in the study. At the pre-test stage, the variance between the groups met the homogeneity criterion. At the post-test stage, however, the variance did not meet the criterion. The study therefore employed a non-parametric statistical test for the subsequent group comparison.

7. Hypothesis Testing

The Mann–Whitney U test was used to compare the experimental and control groups. The test results were reported for both the pre-test and post-test. The pre-test analysis produced a Mann–Whitney U value of 44.500, a Z value of −2.830, and an asymptotic significance value of 0.005. The Mann–Whitney U test showed a statistically significant difference at both pre-test ($p = 0.005$) and post-test ($p < 0.001$).

Table 4. Mann–Whitney U Test Results

| Measurement | Mann–Whitney U | Z | Asymp. Sig. (2-tailed) |
|---|---|---|---|
| Pre-test | 44.500 | −2.830 | 0.005 |
| Post-test | 28.000 | −3.533 | <0.001 |

Source: Authors' data

The post-test significance value of less than 0.001 was lower than the 0.05 significance level. Therefore, the post-test scores of the experimental and control groups differed statistically significantly. This statistical result corresponds with the descriptive findings, in which the experimental group obtained a higher post-test mean score of 81.47 than the control group's mean score of 60.20. The post-test results therefore show a clear statistical difference in writing achievement between the two groups at the end of the study period.

An important result, however, is that the pre-test comparison also produced a significance value below 0.05. The pre-test significance value was 0.005, indicating that the experimental and control groups were already statistically different before the treatment. This result is consistent with the descriptive data, in which the experimental group had a higher pre-test mean score of 50.13 compared with 36.27 in the control group.

The pre-test result is therefore an important characteristic of the data and should be reported transparently. The two groups were not statistically equivalent at baseline. Consequently, the statistically significant post-test difference should be interpreted together with the statistically significant pre-test difference. The post-test result demonstrates that the groups differed significantly after the learning period, while the present analysis does not completely separate the post-test difference from the initial difference between the groups.

The descriptive improvement provides additional information. The experimental group increased by 31.34 points in mean writing score, from 50.13 to 81.47. The control group increased by 23.93 points, from 36.27 to 60.20. Therefore, the experimental group showed a larger descriptive improvement by 7.41 points. Nevertheless, because the groups were already significantly different at pre-test, this descriptive difference should not be interpreted independently as proof that the treatment alone produced the observed post-test difference.

Overall, both groups demonstrated improvement in writing achievement, with a larger descriptive increase in the experimental group (31.34 points) than in the control group (23.93 points). The Shapiro–Wilk test indicated that the pre-test and post-test scores were normally distributed, while Levene's test showed that the post-test variances were not homogeneous. The Mann–Whitney U test revealed statistically significant differences between the groups at both the pre-test ($p = 0.005$) and post-test ($p < 0.001$).

However, the significant pre-test difference indicates that the two groups were not equivalent at baseline. Therefore, although the experimental group achieved a higher post-test mean and a greater descriptive improvement, the findings should be interpreted cautiously and do not provide conclusive evidence that digital journaling alone caused the observed difference.

## Analysis

The analysis of the findings focuses on interpreting the students' writing scores, the improvement observed in both groups, the prerequisite test results, and the hypothesis testing. The interpretation is based on the quantitative data obtained from the pre-test and post-test and is related to previous theoretical and empirical perspectives on writing development, feedback, digital learning, and learner autonomy.

1. Students' Writing Achievement Before and After the Treatment

The descriptive results indicate that both the experimental and control groups demonstrated improvement in writing achievement after the learning process. However, the experimental group showed a higher mean score at both measurement points and a greater descriptive increase from pre-test to post-test. The experimental group obtained a mean score of 50.13 in the pre-test and increased to 81.47 in the post-test. Thus, the mean score increased by 31.34 points. In comparison, the control group increased from 36.27 in the pre-test to 60.20 in the post-test, representing a mean increase of 23.93 points. Therefore, the descriptive difference in mean improvement between the two groups was 7.41 points.

The higher improvement observed in the experimental group can be considered in relation to the characteristics of the digital journaling activities. During the intervention, students were required to write regularly, submit their journal entries through Google Docs, receive lecturer feedback through the commenting feature, revise their writing, and resubmit the revised version. This process provided students with repeated opportunities to practice writing rather than completing writing as a one-time classroom task. The writing process therefore involved drafting, receiving feedback, revising, and producing an improved version.

This pattern is consistent with Hyland's (2022) view that regular writing practice can help learners develop different dimensions of writing, including organization, vocabulary, grammatical accuracy, and overall writing quality. The present data support this perspective at the descriptive level because the experimental group demonstrated a substantial increase in its overall writing scores after participating in repeated digital writing activities. However, the available data do not provide separate scores for content, organization, vocabulary, language use, and mechanics; therefore, the improvement should be interpreted as an overall writing achievement rather than improvement in a particular writing component.

The findings can also be interpreted from the perspective of feedback and revision. Pham (2022) emphasizes the importance of written corrective feedback in helping second-language learners identify and revise problems in their writing. In the present study, feedback was integrated into the digital journaling process through Google Docs. Students could receive lecturer comments and subsequently revise their journal entries. Thus, the greater descriptive improvement in the experimental group may be related to the opportunity to engage in an iterative writing process. The data, however, do not independently measure how much of the improvement was caused by feedback, revision, or the digital platform itself. Therefore, these factors should be understood as components of the treatment rather than as independently tested variables.

The findings are also in line with Sa'diyah and Nabhan (2021) who reported that Google Docs can facilitate collaborative writing, continuous teacher feedback, and revision. Similarly, Kurniawati (2022) found that Google Docs can provide an interactive learning environment that supports communication between lecturers and students during writing activities. In the present study, Google Docs functioned not only as a place to submit writing but also as a medium through which students received feedback and revised their work. This distinction is important because the intervention was not simply the use of digital technology; it involved the integration of digital technology into a continuous writing process (Tian, Liu, & Zhang, 2022).

2. Comparison of Improvement Between the Experimental and Control Groups

The descriptive comparison provides evidence of a different pattern of improvement between the two groups. The experimental group increased from 50.13 to 81.47, whereas the control group increased from 36.27 to 60.20. The experimental group therefore achieved both a higher post-test mean and a larger mean increase. In addition, all students in the experimental group demonstrated higher post-test scores than their corresponding pre-test scores. In the control group, most students improved, although two students maintained the same score in the pre-test and post-test.

This pattern suggests that both conventional instruction and digital journaling were accompanied by improvement in students' writing performance. Consequently, the improvement cannot be interpreted as occurring only in the experimental group. The difference is that the experimental group showed a larger descriptive increase. This is an important finding because it indicates that digital journaling may provide additional learning opportunities beyond conventional classroom writing instruction.

The use of digital journaling may have supported the continuity of writing practice. Students were not limited to producing writing only during scheduled classroom activities. They could access and revise their journal entries through Google Docs, while lecturer feedback could be provided through the digital commenting feature. Iftanti (2024) emphasizes the importance of digital literacy for university students in the context of Society 5.0. The present findings are compatible with this perspective because the experimental treatment required students to engage repeatedly with their own written production through a digital platform. The present findings are compatible with this perspective because the experimental treatment required students to engage repeatedly with their own written production.

Nevertheless, the data should not be interpreted as demonstrating that digital journaling was the only factor responsible for the improvement. The experimental group and control group followed different instructional conditions, and the study did not separately measure students' engagement, amount of writing, number of revisions, or frequency of feedback. Therefore, the observed improvement should be interpreted within the overall treatment condition.

The findings are also consistent with studies by Sembiring et al. (2025) which reported positive applications of Google Docs in EFL writing contexts. They found that Google Docs could support interaction, peer support, and participation in collaborative writing. They reported that Google Docs-based collaborative writing activities could improve students' engagement and writing quality. The present study extends this line of research to a traditional Islamic boarding school involving university students. The contextual contribution is therefore important because digital writing practices were implemented in an educational environment that combines academic and Islamic learning experiences

3. Interpretation of the Normality and Homogeneity Tests

The Shapiro–Wilk test showed significance values of 0.664 for the experimental pre-test, 0.660 for the control pre-test, 0.078 for the experimental post-test, and 0.140 for the control post-test. All values were higher than 0.05. Therefore, based on the reported criterion, the four sets of scores were normally distributed.

The Levene's test produced a significance value of 0.061 for the pre-test and 0.011 for the post-test. The pre-test therefore met the reported homogeneity criterion, whereas the post-test did not because its significance value was below 0.05.

These prerequisite tests provide the statistical basis for the analysis reported in the study. Because the post-test data did not meet the homogeneity assumption, the study used the Mann–Whitney U test as the non-parametric comparison reported in the article. The statistical analysis therefore considered not only the descriptive differences but also the distributional characteristics of the data.

4. Interpretation of the Hypothesis Testing

The Mann–Whitney U test produced a significance value of 0.005 for the pre-test and less than 0.001 for the post-test. The post-test result indicates a statistically significant difference between the experimental and control groups after the learning period because the significance value was below the 0.05 criterion. The experimental group also had a substantially higher post-test mean score than the control group, with values of 81.47 and 60.20, respectively.

However, an important statistical finding must be considered: the pre-test comparison was also statistically significant, with $p = 0.005$. This means that the two groups were already significantly different before the treatment. The descriptive data support this result because the experimental group started with a higher pre-test mean of 50.13, whereas the control group had a pre-test mean of 36.27.

This baseline difference has an important consequence for interpreting the effectiveness of the treatment. The statistically significant post-test difference demonstrates that the two groups differed significantly after the learning period, but the current analysis cannot completely separate the post-test difference from the difference that already existed before the treatment. Therefore, the result should not be interpreted as definitive evidence that digital journaling alone caused the statistically significant post-test difference.

At the same time, the descriptive results remain meaningful. The experimental group increased by 31.34 points, compared with 23.93 points in the control group. This indicates that the experimental group demonstrated a greater descriptive improvement during the study. The finding is therefore supportive of the potential value of digital journaling, but the strength of the causal claim should be limited because the groups were not equivalent at baseline.

From a research perspective, this finding highlights the importance of considering initial group differences in quasi-experimental research. The non-equivalent control group design does not guarantee that participants in the two groups have the same initial level of ability. Therefore, baseline differences need to be taken into account when interpreting post-test comparisons. The current findings should consequently be understood as evidence of a significant post-test difference accompanied by a greater descriptive increase in the experimental group, rather than as conclusive proof of a treatment-only effect.

5. Theoretical and Practical Interpretation

The findings provide several implications for English writing instruction. First, the increase in the experimental group is consistent with the view that writing development

requires continuous practice. Hyland (2022) emphasizes writing as a process that involves the development and organization of ideas rather than simply producing a final written product. Digital journaling provides a context in which students can repeatedly engage in writing and revision.

Second, the findings support the role of feedback in the writing process. Through Google Docs, students could receive lecturer comments and revise their writing. This is compatible emphasis on feedback and revision in second-language writing. The digital environment may make the feedback process more accessible because comments can be attached directly to students' written work.

Third, the findings have implications for digital literacy and learner autonomy. Iftanti et al. (2024) emphasize the increasing importance of digital literacy for university students in the Society 5.0 era. In the present treatment, students used Google Docs to prepare, revise, and submit their writing. This activity therefore combined language practice with the use of digital tools. However, because digital literacy and learner autonomy were not measured using separate instruments in this study, the findings should be interpreted as potential instructional implications rather than direct empirical evidence of improvement in those constructs.

Overall, the analysis indicates that digital journaling through Google Docs was associated with higher writing achievement at the end of the learning period and a larger descriptive increase than conventional writing instruction. The finding is consistent with previous research on journal writing, feedback, revision, and Google Docs-supported writing. Nevertheless, the significant pre-test difference requires a cautious interpretation of the causal effect. The principal contribution of the study is therefore the empirical observation that digital journaling can be implemented as a continuous writing activity in an Islamic boarding school context and is accompanied by substantial improvement in students' writing scores.

## Conclusion

This study examined the implementation of digital journaling activities through Google Docs in improving university students' English writing achievement at Subulussalam Islamic Boarding School. The findings showed that both the experimental and control groups improved their writing performance after the learning process. The experimental group increased from a mean pre-test score of 50.13 to 81.47 in the post-test, representing a 31.34-point increase, while the control group increased from 36.27 to 60.20, representing a 23.93-point increase. Thus, the experimental group demonstrated a greater descriptive improvement than the control group. The Mann–Whitney U test also showed a statistically significant difference between the groups in the post-test ($p < 0.001$).

However, the pre-test comparison was also statistically significant ($p = 0.005$), indicating that the two groups were not equivalent at the beginning of the study. Therefore, the post-test difference should be interpreted cautiously and should not be considered conclusive evidence that digital journaling alone caused the observed difference. Nevertheless, the higher post-test performance and greater descriptive improvement of the experimental group indicate the potential value of digital journaling through Google Docs as a writing-learning activity.

The study contributes to English language education by illustrating how digital journaling can provide continuous opportunities for writing practice, feedback, and revision in an Islamic boarding school context. Practically, English teachers and lecturers may consider integrating Google Docs into writing instruction to support regular writing

activities and digital learning. Future research should involve larger samples, longer intervention periods, and statistical procedures that specifically control for initial differences between groups to provide stronger evidence regarding the effectiveness of digital journaling.

## References

Alwahoub, H. M., Jomaa, N. J., & Azmi, M. N. L. (2022). The impact of synchronous collaborative writing and Google Docs collaborative features on enhancing students' individual writing performance. Indonesian Journal of Applied Linguistics, 12(1), 111–123. https://doi.org/10.17509/ijal.v12i1.46541

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

Hyland, K. (2022). Teaching and researching writing (4th ed.). Routledge.

Iftanti, E., Madayani, N. S., & Ningrum, A. S. B. (2024). Digital literacy development of undergraduate students for Society 5.0: Pedagogical implication for language and science education. Journal of English Education and Linguistics Studies, 12(1), 109–140. https://doi.org/10.30762/jeels.v12i1.3813

Kurniawati, A. (2022). Google Docs to manage an EFL writing class: How it helps and what to prepare. Journal of English and Education, 8(2), 97–109. https://doi.org/10.20885/jee.v8i2.25740

Nur, A. W., & Ramadhani, L. F. (2025). The use of digital learning media in EFL classroom for writing skills. JEELL: Journal of English Education, Linguistics and Literature, 12(2), 92–110. https://doi.org/10.32682/jeell.v12i2.60

Pham, H. T. P. (2022). Computer-mediated and face-to-face peer feedback: Student feedback and revision in EFL writing. Computer Assisted Language Learning, 35(9), 2112–2147. https://doi.org/10.1080/09588221.2020.1868530

Phuong, H. Y., Phan, Q. T., & Le, T. T. (2023). The effects of using analytical rubrics in peer and self-assessment on EFL students' writing proficiency: A Vietnamese contextual study. Language Testing in Asia, 13, 42. https://doi.org/10.1186/s40468-023-00256-y

Sa'diyah, H., & Nabhan, S. (2021). Collaborative writing using Google Docs in an EFL classroom: Voices from high school students. VELES: Voices of English Language Education Society, 5(2), 156–166. https://doi.org/10.29408/veles.v5i2.3863

Sembiring, S. S., Yuliati, & Sakhiyya, Z. (2025). Exploring the use of Google Docs for collaborative narrative writing: A sequential exploratory study in EFL classroom. ETERNAL: English Teaching Journal, 16(2), 442–456. https://doi.org/10.26877/eternal.v16i2.1543

Tian, L., Liu, Q., & Zhang, X. (2022). Self-regulated writing strategy use when revising upon automated, peer, and teacher feedback in an online English as a foreign language writing course. Frontiers in Psychology, 13, 873170. https://doi.org/10.3389/fpsyg.2022.873170

Valizadeh, M. (2022). Collaborative writing on Google Docs: Effects on EFL learners' descriptive paragraphs. Indonesian Journal of English Language Teaching and Applied Linguistics, 6(2), 277–287. https://doi.org/10.21093/ijeltal.v6i2.1053

Zahrok, A., & Rahmawati, E. (2025). The impact of online collaborative writing strategy via Google Docs on senior high school learners' writing performance in composing narrative texts. LET: Linguistics, Literature and English Teaching Journal, 15(2), 356–384. https://doi.org/10.18592/let.v15i2.17742