

The Utilization of Artificial Intelligence (Ai) Writing Assistants to the Writing Proficiency

Rhona Mae M. Coracero

julierosemendoza002@gmail.com

Laguna State Polytechnic University Sta. Cruz Laguna 4009 PHILIPPINES

Abstract

This study aimed to determine the utilization of AI writing assistants to the writing proficiency of senior high school students. Specifically, it aimed to identify the status of utilization of AI writing assistants with regards to grammar and spelling correction, sentence structuring, word choice and vocabulary suggestion, summarization, and paraphrasing; the level of the writing proficiency of the Grade 11 senior high school students in their diagnostic, formative, and summative tests; the significant difference in the writing proficiency of the Grade 11 students in their diagnostic, formative, and summative tests as well as in between groups; and the significant effect of the utilization of AI writing assistants to the writing proficiency of the Grade 11 senior high school students of Cristobal S. Conducto Memorial Integrated National High School (CSCMINHS).

The study was conducted at CSCMINHS in Rizal, Laguna, from October 2023 to March 2024. The researcher used an experimental research design with two experimental groups and a control group, involving 107 Grade 11 students enrolled in the General Academic Strand (GAS). Data were gathered through the use of researcher-made writing proficiency test and questionnaire.

Based on the results, the level of utilization of AI writing assistants is high. In terms of writing proficiency, respondents demonstrated fairly satisfactory to very satisfactory level of skills. Moreover, significant differences were observed in the diagnostic, formative and summative test results of both the experimental and control groups, indicating changes in writing proficiency over the assessment periods. However, out of the five indicators of writing proficiency, only two areas—grammar usage and critical thinking skills were found to have significant differences between the experimental and control groups. Since most writing proficiency areas did not show any significant difference, the researcher concludes that there is no difference in the writing proficiency of the experimental and control groups.

Overall, despite the observed differences in two indicators of writing proficiency, the study found no significant effect of AI writing assistants on the writing proficiency of senior high school students. This led to the acceptance of the null hypothesis made by the researcher at the beginning of the study.

The recommendations stemming from these findings include further studies on writing proficiency focusing on grammar usage and critical thinking, investigation on the integration of AI writing assistants in writing activities in class, and the conduct of longitudinal study to accurately assess long-term effect of AI writing assistants to the writing proficiency of students.

Keywords: writing proficiency; utilization of AI; writing assistants

1. Introduction

Writing is an essential part of communication and is one of the four macro skills developed in schools. Writing proficiency allows one to deliver the message clearly and effectively. It is an arduous process that challenges a person's critical thinking skills and knowledge of spelling, grammar, vocabulary, organization, and others (Ravichandran et al., 2023). In general, mastering the skill of writing is not an easy task, but it is far more challenging to master writing in one's second language (L2) (Fareed et al., 2016).

Many writing assistants on the web are utilized to generate ideas, organize thoughts, and produce content more efficiently. These writing assistants are powered by artificial intelligence (AI), which mimics human writing and can produce real-time feedback on its users' writing. Due to this, students have been seeking the help of AI writing assistants to aid them in their writing tasks or activities by correcting errors in their content.

While AI writing assistants can aid teachers in grading and providing feedback on students' output, they also threaten their jobs. Students no longer seek their comments and advice due to the accessibility of AI writing assistants and its ability to provide feedback immediately. However, students who become dependent on these AI writing assistants may potentially have decreased critical thinking and problem solving skills (Moonpreneur, 2023). Since AI writing assistants can generate content, there is also an issue with the ownership and accountability of the written content. Students may submit AI-generated content and declare it as their writing.

In connection to this, the study is conducted to identify the role of AI writing assistants to the writing proficiency of senior high school students. It also aims to identify the possible implication of the use of AI writing assistants as a pedagogical tool to improve students' writing. Moreover, the result of this study can possibly pinpoint areas for improvement in the students' writing such as grammar, spelling, syntax, and others.

1.1 Statement of the Problem

Specifically, this sought to answer the following questions:

1. What is the status of utilization of AI writing assistants with regards to:
 - 1.1. Grammar and Spelling Correction;
 - 1.2. Sentence Structuring;
 - 1.3. Word Choice and Vocabulary Suggestion; and
 - 1.4. Summarization and Paraphrasing?
2. What is the level of the writing proficiency of the Grade 11 senior high school students in their Diagnostic, Formative, and Summative tests in terms of:
 - 2.1. Grammar;
 - 2.2. Mechanics;
 - 2.3. Organization;
 - 2.4. Vocabulary; and
 - 2.5. Critical Thinking?
3. Is there a significant difference in the writing proficiency of the Grade 11 SHS students in their:
 - 3.1. Diagnostic and Formative Test; and
 - 3.2. Formative and Summative Test?
4. Is there a significant difference in the writing proficiency of the experimental group and control group?

5. Does the utilization of AI writing assistants have significant effect to the writing proficiency of the Grade 11 Senior High School students of Cristobal S. Conducto Memorial Integrated National High School?

2. Methodology

This study used quantitative research design to identify the utilization of AI writing assistants to the writing proficiency of senior high school students. Specifically, this study used experimental research design, which refers to the testing and comparing of two or more variables in controlled environments (Sirisilla & Sirisilla, 2023).

In this study, there were two experimental groups in which the effects of AI writing assistants were tested, and a control group to which the result of the former will be compared.

3. Results and Discussion

This chapter presents, analyzes, and interprets the data gathered from the respondents by providing answers to the questions on the status of utilization of AI writing assistants in terms of grammar and spelling correction, sentence structuring, word choice and vocabulary suggestion, and summarization and paraphrasing. Moreover, this chapter also contains answer to questions on the level of the students' writing proficiency in terms of grammar usage, mechanics, organization, vocabulary, and critical thinking. In addition, the chapter also presents data gathered pertaining to the difference in writing proficiency of the experimental and controlled group as well as the effect of the utilization of AI writing assistance on the writing proficiency of the students.

Status of Utilization of AI Writing Assistants

In this study, the independent variable measured is the status of utilization of AI writing assistants. As the use of such tools had been rampant in the academic sector, the researcher investigated the extent to which students use AI writing assistants in their writing. Indicators of the utilization of AI writing assistants includes grammar and spelling correction, sentence structuring, word choice and vocabulary suggestion, and summarization and paraphrasing.

To determine the level of the status of utilization of AI writing assistants, the researcher used mean and standard deviation in interpreting the data gathered from the respondents.

Table 1. Status of Utilization of AI Writing Assistants with regards to Grammar and Spelling Correction

Indicators	Mean	SD	Remarks
<i>Artificial Intelligence writing assistants...</i>			
Help me catch and correct errors in grammar that I might overlook.	4.27	0.67	Strongly Agree
Have improved my understanding of proper grammar rules.	4.02	0.64	Agree
Provide clear explanations for suggested grammar and spelling corrections.	4.14	0.77	Agree
Accurately detect and suggest corrections for misspelled words.	4.11	0.75	Agree

Distinguish effectively homophones and suggests corrections.	3.93	0.73	Agree
--	------	------	-------

Overall Mean = 4.09

Standard Deviation = 0.72

Verbal Interpretation = High

Table 1 presents the results of a survey on the status of utilization of AI writing assistants in the context of grammar and spelling correction. Respondents strongly agree that AI writing assistants help them catch and correct errors in grammar that they might overlook ($M = 4.27$). This indicates a high level of confidence in the ability of these assistants to assist with grammar correction. While respondents agree that AI writing assistants effectively distinguish homophones and suggest corrections ($M = 3.93$). The overall mean of all indicators is 4.09 with the standard deviation of 0.72 indicating a generally high level of satisfaction with the performance of AI writing assistants in grammar and spelling correction. The results imply that AI writing assistants are widely regarded as effective tools for grammar and spelling correction, offering educational value, accuracy, and user satisfaction. However, there may still be room for refinement in specific areas such as handling homophones to further enhance their utility and performance.

Table 2. Status of Utilization of AI writing Assistants with regards to Sentence Structuring

Indicators	Mean	SD	Remarks
<i>Artificial Intelligence writing assistants...</i>			
Increase text readability.	3.95	0.69	Agree
Assist me in avoiding repetitive sentence structures.	3.79	0.70	Agree
Aid in writing coherent and cohesive texts, with organized and well-connected ideas.	3.86	0.77	Agree
Suggest appropriate sentence constructions for my writing.	4.07	0.68	Agree
Provide helpful suggestions for sentence length and complexity.	4.20	0.79	Agree

Overall Mean = 3.97

Standard Deviation = 0.74

Verbal Interpretation = High

Table 2 presents the results of a survey on the status of utilization of AI writing assistants with regards to sentences structuring. Respondents agree that AI writing assistants provide helpful suggestions for sentence length and complexity ($M = 4.20$). This indicates a good level of confidence in the ability of these assistants to provide feedback with regards to sentences' structure. Respondents also agree that AI writing assistants assist them in avoiding repetitive sentence structures ($M = 3.79$). The overall mean of all indicators is 3.97 with the standard deviation of 0.74 indicating a generally high level of satisfaction with the performance of AI writing assistants with regards to sentence structuring. The results imply that AI writing assistants are generally considered as effective tools for sentence structuring. However, there may still be room for refinement in specific areas such as assisting in avoiding repetitive structures in sentences.

Table 3. Status of Utilization of AI writing Assistants with regards to Word Choice and Vocabulary

Suggestion

Indicators	Mean	SD	Remarks
Artificial Intelligence writing assistants...			
Are valuable tools for refining the word choices and vocabulary in my documents.	3.93	0.68	Agree
Help me avoid repetitive words and phrases in my writing.	3.91	0.61	Agree
Offer suggestions for improved readability of the text.	3.84	0.80	Agree
Effectively identify and suggest more precise or appropriate words for my content.	4.04	0.73	Agree
Offer valuable alternatives for enhancing the vocabulary in my writing.	3.96	0.84	Agree

Overall Mean = 3.94

Standard Deviation = 0.74

Verbal Interpretation = High

Table 3 presents the results of a survey on the status of utilization of AI writing assistants with regards to word choice and vocabulary suggestion. Respondents agree that AI writing assistants effectively identify and suggest more precise or appropriate words for the content (M= 4.04). This indicates a good level of confidence in the ability of these assistants to present more appropriate words for the text being written. Respondents also agree that AI writing assistants offer suggestions for improved readability of the text (M=3.84). The overall mean of all indicators is 3.94 with the standard deviation of 0.74 indicating a generally high level of satisfaction with the performance of AI writing assistants with regards to word choice and vocabulary suggestion. The results suggest that AI writing assistants are widely viewed as effective tools for word choice and vocabulary suggestion. However, there may still be room for improvement in specific areas such as offering suggestions for improved readability of the text, avoiding repetitive words and phrases in writing, and refining word choices and vocabulary in document.

Table 4. Status of Utilization of AI writing Assistants with regards to Summarization and Paraphrasing

Indicators	Mean	SD	Remarks
Artificial Intelligence writing assistants...			
Effectively summarize text while retaining key information.	4.07	0.78	Agree
Have improved my ability to create clear and short summaries.	4.07	0.78	Agree
Effectively paraphrase sentences in my writing.	3.82	0.80	Agree
Help me avoid unintentional plagiarism through effective paraphrasing.	3.86	0.85	Agree
Offer diverse options for paraphrasing, allowing me to choose the	4.09	0.74	Agree

 most suitable alternatives.

Overall Mean = 3.98

Standard Deviation = 0.80

Verbal Interpretation = High

Table 4 presents the results of a survey on the status of utilization of AI writing assistants with regards to summarization and paraphrasing. Respondents agree that AI writing assistants offer diverse options for paraphrasing, allowing one to choose the most suitable alternatives ($M=4.09$). This indicates a good level of confidence in the ability of these assistants to generate a suitable paraphrased text of the inputted material. Respondents also agree that AI writing assistants effectively paraphrase sentences in writing ($M=3.82$). The overall mean of all indicators is 3.98 with the standard deviation of 0.80 indicating a generally good level of satisfaction with the performance of AI writing assistants with regards to summarization and paraphrasing. The results imply that AI writing assistants are widely regarded as effective tools for summarization and paraphrasing. However, there may still be room for refinement in specific areas such as effectively paraphrasing sentences in writing.

Level of the Writing Proficiency

The level of writing proficiency of the students comprises grammar, mechanics, organization, vocabulary, and critical thinking was measured by getting the mean and standard deviation.

Table 5. Level of the Writing Proficiency of the Grade 11 Senior High School students in terms of Grammar Usage

Indicators	Group A			Group B			Group C		
	Mean	SD	Remarks	Mean	SD	Remarks	Mean	SD	Remarks
Diagnostic Test	4.68	1.73	Satisfactory	4.39	1.23	Fairly Satisfactory	5.30	1.47	Satisfactory
Formative Test	5.26	1.39	Satisfactory	4.64	1.42	Satisfactory	6.39	1.30	Satisfactory
Summative Test	6.29	1.47	Satisfactory	5.03	1.11	Satisfactory	6.39	1.39	Satisfactory

Table 5 shows the level of writing proficiency of Grade 11 Senior High School students across three different groups such as experimental group with the intervention of AI writing assistants (group A), another experimental with half the group with intervention of AI writing assistants (group B), and no intervention of AI writing assistants or the controlled group (group C) measured through diagnostic, formative, and summative tests, with focus on grammar proficiency.

In the diagnostic test, Group A attained satisfactory level of proficiency ($M=4.68$), while group B attained slightly lower than Group A but fairly satisfactory level of proficiency ($M=4.39$). The highest level of proficiency was attained by group C as compared to Groups A and B ($M=5.30$).

In the formative test, Group A attained satisfactory level of proficiency, similar to the diagnostic test. Some variability in scores is observed, but the mean remains at a satisfactory level. ($M=5.26$). Group B also has satisfactory level of proficiency, but slightly lower than Group A ($M=4.64$). Group C has the highest level of proficiency as compared to the other groups. Both the mean score indicates a satisfactory level of

proficiency performing notably better on average ($M=6.39$)

In the summative test, Group A attained satisfactory level of proficiency, similar to the diagnostic test ($M= 6.29$). Group B gained satisfactory level of proficiency, slightly lower than group A. Group C has a significantly higher level of proficiency as compared to the other groups ($M= 6.39$).

Overall, the interpretation means that group C consistently outperforms the other groups across all types of tests, demonstrating the highest level of writing proficiency in terms of grammar. Moreover, Group A and Group B showed improvement over time. Additionally, the variability in scores, as indicated by the standard deviations, suggests differences in the distribution of proficiency levels within each group.

Table 6. Level of the Writing Proficiency of the Grade 11 Senior High school Students in terms of Mechanics

Indicators	Group A			Group B			Group C		
	Mean	SD	Remarks	Mean	SD	Remarks	Mean	SD	Remarks
Diagnostic Test	3.84	2.83	Fairly Satisfactory	4.56	2.14	Fairly Satisfactory	6.24	62.14	Satisfactory
Formative Test	5.97	2.01	Satisfactory	5.56	2.09	Satisfactory	6.36	1.78	Satisfactory
Summative Test	6.26	1.73	Satisfactory	6.53	1.66	Very Satisfactory	6.82	1.74	Very Satisfactory

Table 6 shows the level of writing proficiency of Grade 11 Senior High School students across three different groups such as experimental group with the intervention of AI in writing (group A), another experimental with half the group with intervention of AI writing (group B), and no intervention of AI writing or the controlled group (group C) measured through diagnostic, formative, and summative tests, with focus on mechanics.

In the diagnostic test, Group A and Group B attained fairly satisfactory level of proficiency ($M=3.84$; $M=4.56$, respectively), while group C attained satisfactory level of proficiency ($M=6.24$).

In the formative test, Group A attained satisfactory level of proficiency ($M=5.97$), significantly higher than the diagnostic test. Group B also attained satisfactory level of proficiency ($M=5.56$), which is also higher than the diagnostic test. Group C showed some variability in scores, but the mean remains at a satisfactory level ($M=6.36$).

In the summative test, Group A attained satisfactory level of proficiency, similar to the formative test ($M= 6.26$). Group B gained very satisfactory level of proficiency, higher than group A. Group C also has a significantly higher level of proficiency as compared to the other groups ($M= 6.82$).

Overall, the interpretation means that group C consistently outperforms the other groups across all types of tests, demonstrating the highest level of writing proficiency in terms of mechanics. Group A shows improvement over time, while group B demonstrates more progress in terms of mechanics in writing. Additionally, the variability in scores, as indicated by the standard deviations, suggests differences in the distribution of proficiency levels within each group.

Table 7. Level of the writing proficiency of the Grade 11 Senior High School students in terms of Organization

Indicators	Group A			Group B			Group C		
	Mean	SD	Remarks	Mean	SD	Remarks	Mean	SD	Remarks
Diagnostic Test	4.53	1.75	Fairly Satisfactory	4.31	1.60	Fairly Satisfactory	5.06	1.64	Satisfactory
Formative Test	5.47	1.66	Satisfactory	4.89	1.67	Satisfactory	5.21	1.71	Satisfactory
Summative Test	6.05	1.56	Satisfactory	5.25	1.68	Satisfactory	5.64	1.50	Satisfactory

Table 7 shows the level of writing proficiency of Grade 11 Senior High School students across three different groups such as experimental group with the intervention of AI in writing (group A), another experimental with half the group with intervention of AI writing (group B), and no intervention of AI writing or the controlled group (group C) measured through diagnostic, formative, and summative tests, with focus on organization.

In the diagnostic test, Group A attained fairly satisfactory level of proficiency ($M=4.53$), while group B attained slightly lower than Group A but fairly satisfactory level of proficiency ($M=4.31$). The highest level of proficiency was attained by group C as compared to Groups A and B ($M= 5.06$).

In the formative test, Group A gained satisfactory level of proficiency. Group B also has satisfactory level of proficiency, but slightly lower than Group A ($M=4.89$). Group C has the highest level of proficiency as compared to the other groups. Both the mean score indicates a satisfactory level of proficiency performing notably better on average ($M=5.21$)

In the summative test, Group A attained the highest satisfactory level of proficiency as compared to the other groups ($M= 6.05$). Group B gained satisfactory level of proficiency, slightly lower than group A. Group C has also has satisfactory level of proficiency ($M= 5.64$).

Overall, the interpretation means that group A has the biggest improvement over time, while group B demonstrates modest gains. Group C has minimal changes in scores. Additionally, the variability in scores, as indicated by the standard deviations, suggests differences in the distribution of proficiency levels within each group.

Table 8. Level of the Writing Proficiency of the Grade 11 Senior High School Students in terms of Vocabulary

Indicators	Group A			Group B			Group C		
	Mean	SD	Remarks	Mean	SD	Remarks	Mean	SD	Remarks
Diagnostic Test	3.71	2.17	Fairly Satisfactory	4.86	1.85	Satisfactory	4.91	1.79	Satisfactory
Formative Test	4.58	2.33	Fairly Satisfactory	5.28	1.92	Satisfactory	5.70	1.91	Satisfactory
Summative Test	5.58	1.91	Satisfactory	6.03	1.98	Satisfactory	5.73	1.94	Satisfactory

Table 8 shows the level of writing proficiency of Grade 11 Senior High School students across three different groups such as experimental group with the intervention of AI in writing (group A), another experimental with half the group with intervention of AI writing (group B), and no intervention of AI writing

or the controlled group (group C) measured through diagnostic, formative, and summative tests, with focus on vocabulary.

In the diagnostic test, Group A attained fairly satisfactory level of proficiency ($M=3.71$), while group B attained a higher result than Group A with satisfactory level of proficiency ($M=4.86$). The highest level of proficiency was attained by group C as compared to Groups A and B ($M= 4.91$).

In the formative test, Group A attained fairly satisfactory level of proficiency, similar to the diagnostic test. Some variability in scores is observed, but the mean remains at a satisfactory level ($M=4.58$). Group B has satisfactory level of proficiency ($M=5.28$). Group C has the highest level of proficiency as compared to the other groups. Both the mean score indicates a satisfactory level of proficiency performing notably better on average ($M=5.70$)

In the summative test, Group A attained satisfactory level of proficiency ($M= 5.58$). Group B also gained satisfactory level of proficiency, slightly higher than group A ($M=6.03$). Group C has satisfactory level of proficiency ($M=5.73$).

Overall, the interpretation means that group C outperforms the other groups across all types of tests, except on the summative test on vocabulary. Group A shows improvement over time, while group B demonstrates more modest gains. Additionally, the variability in scores, as indicated by the standard deviations, suggests differences in the distribution of proficiency levels within each group.

Table 9. Level of the Writing Proficiency of the Grade 11 Senior High School Students in terms of Critical Thinking

Indicators	Group A			Group B			Group C		
	Mean	SD	Remarks	Mean	SD	Remarks	Mean	SD	Remarks
Diagnostic Test	4.18	2.26	Fairly Satisfactory	4.47	2.30	Fairly Satisfactory	4.85	2.21	Satisfactory
Formative Test	5.21	2.20	Satisfactory	5.75	2.45	Satisfactory	5.42	1.44	Satisfactory
Summative Test	6.58	2.07	Very Satisfactory	7.36	1.69	Very Satisfactory	6.21	1.71	Satisfactory

Table 9 shows the level of writing proficiency of Grade 11 Senior High School students across three different groups such as experimental group with the intervention of AI in writing (group A), another experimental with half the group with intervention of AI writing (group B), and no intervention of AI writing or the controlled group (group C) measured through diagnostic, formative, and summative tests, with focus on critical thinking.

In the diagnostic test, Group A attained fairly satisfactory level of proficiency ($M=4.18$), while group B attained slightly higher than Group A but fairly satisfactory level of proficiency ($M=4.47$). The highest level of proficiency was attained by group C as compared to Groups A and B ($M= 4.85$).

In the formative test, Group A attained satisfactory level of proficiency. Group B also has satisfactory level of proficiency, but slightly higher than Group A ($M=5.75$). Group C has also satisfactory level of proficiency ($M=5.42$).

In the summative test, Group A attained very satisfactory level of proficiency ($M= 6.58$). Group B also gained very satisfactory level of proficiency, slightly higher than group A ($M=7.36$). Group C has the lowest level of proficiency as compared to the other groups ($M= 6.21$).

Overall, the interpretation means that group A and B show great improvement over time, while group C demonstrates small gains. Additionally, the variability in scores, as indicated by the standard deviations,

suggests differences in the distribution of proficiency levels within each group.

Significant Difference in the Writing Proficiency

The significant difference in writing proficiency of grade 11 SHS students includes the comparison between diagnostic and formative test, and between formative and summative test.

Table 10. Significant Difference in the Writing Proficiency of the Grade 11 SHS students between Diagnostic and Formative Test

Writing Proficiency	Diagnostic		Formative		Mean Difference	<i>t</i>	<i>df</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Grammar	4.78	1.53	5.40	1.53	0.63	-2.99	212	0.003
Mechanics	4.82	2.59	5.95	1.98	1.13	-3.59	212	0.000
Organization	4.62	1.68	5.20	1.68	0.58	-2.52	212	0.012
Vocabulary	4.47	2.02	5.16	2.11	0.69	-2.45	212	0.015
Critical Thinking	4.49	2.25	5.46	2.08	0.97	-3.27	212	0.001

Note: * $p < .05$ significant

Table 10 presents the comparison of writing proficiency scores among Grade 11 Senior High School students between Diagnostic and Formative Tests across different components such as grammar usage, mechanics, organization, vocabulary, and critical thinking.

There are significant differences between the writing proficiency scores of students in Diagnostic and Formative Tests, as indicated by the ($p=0.003$) grammar, ($p=0.0000$) mechanics, ($p=0.012$) organization, ($p=0.015$) vocabulary and ($p=0.001$) critical thinking being less than 0.05 level of significance. The mean differences in scores between the Formative and Diagnostic Tests are positive for all components, indicating that, students scored higher in the Formative Test compared to the Diagnostic Test, *t*-values are negative, this means that the Formative Test scores are significantly lower than the Diagnostic Test scores, which is consistent with the negative mean differences.

Significant Difference in the Writing Proficiency of the Grade 11 SHS students between Formative Test and Summative Test

The significant difference in the writing proficiency of the Grade 11 SHS students between Formative Test and Summative Test was evaluated using the *t*-test assuming equal variances.

Table 11. Significant Difference in the Writing Proficiency of the Grade 11 SHS students between Formative Test and Summative Test

Writing Proficiency	Formative		Summative		Mean Difference	<i>t</i>	<i>df</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Grammar	5.40	1.53	5.90	1.46	0.50	-2.42	212	0.016
Mechanics	5.95	1.98	6.52	1.71	0.57	-2.25	212	0.025
Organization	5.20	1.68	5.65	1.60	0.46	-2.04	212	0.043
Vocabulary	5.16	2.11	5.78	1.93	0.62	-2.23	212	0.027
Critical Thinking	5.46	2.08	6.73	1.89	1.27	-4.68	212	0.000

Note: * $p < .05$. significant

Table 11 presents the comparison of writing proficiency scores among Grade 11 Senior High School students between Formative and Summative Tests across different components such as grammar usage, mechanics, organization, vocabulary, and critical thinking.

There are significant differences between the writing proficiency scores of students in Formative and Summative Tests, as indicated by the ($p=0.016$) grammar, ($p= 0.025$) mechanics, ($p= 0.043$) organization, ($p= 0.027$) vocabulary and ($p= 0.000$) critical thinking being less than 0.05 level of significance. The mean differences in scores between the Formative and Summative Tests are positive for all components, indicating that, students scored higher in the Summative Test compared to the Formative Test, t -values are negative, this means that the Summative Test scores are significantly lower than the Formative Test scores, which is consistent with the negative mean differences.

Significant Difference in the Writing Proficiency of the Experimental Group and Control Group

The significant difference in the writing proficiency of the experimental group and control group was evaluated using the Analysis of Variance (ANOVA). Differences in the areas of grammar usage, mechanics, organization, vocabulary, and critical thinking between the control group and the experimental group are presented in the table. To determine the significant differences, the researcher used Analysis of Variance (ANOVA).

Table 12. Significant Difference in the Writing Proficiency of the Experimental Group and Control Group

Writing Proficiency	Source of Variation	SS	df	MS	F	P-value	F crit
Grammar Usage	Between Groups	22.66	2.00	11.33	3.70	0.03	3.08
	Within Groups	318.50	104	3.06			
	Total	341.16	106				
Mechanics	Between Groups	9.84	2	4.92	1.25	0.29	3.08
	Within Groups	408.07	104	3.92			
	Total	417.91	106				
Organization	Between Groups	11.60	2	5.80	1.69	0.19	3.08
	Within Groups	357.03	104	3.43			
	Total	368.64	106				
Vocabulary	Between Groups	10.70	2	5.35	1.23	0.30	3.08
	Within Groups	453.73	104	4.36			
	Total	464.43	106				
Critical Thinking	Between Groups	24.06	2	12.03	3.54	0.03	3.08
	Within Groups	353.08	104	3.40			
	Total	377.14	106				

Note: * $p < .05$ significant

Table 12 depicts the results of an analysis comparing the writing proficiency of an experimental group with a control group across different aspects such as grammar, mechanics, organization, vocabulary, and critical thinking. There is a significant difference between the experimental and control groups in terms of grammar and critical thinking skill as indicated by the p -values of 0.03. This means that the experimental intervention likely had an effect on the aspects of writing proficiency compared to the control group.

However, mechanics, organization, and vocabulary, is no significant differences between the groups, as indicated by the p-values greater than 0.05. This implies that the experimental intervention did not statistically significant on writing proficiency compared to the control group.

Significant Effect of Utilization of AI Writing Assistants to the Writing Proficiency of the Grade 11 Senior High School students

The significant effect of utilization of AI writing assistants to the writing proficiency of the Grade 11 senior high school students was evaluated using regression analysis.

Table 13. Significant Effect of Utilization of AI Writing Assistants to the Writing Proficiency of the Grade 11 Senior High School students

Writing Proficiency	Beta	SE	95 % CI		β	p
			LL	UL		
Grammar	-0.172	0.508	-1.19	0.845	-0.046	0.735
Mechanics	-1.084	0.596	-2.278	0.11	-0.24	0.074
Organization	0.264	0.536	-0.81	1.338	0.067	0.624
Vocabulary	-0.49	0.7	-1.894	0.914	-0.095	0.487
Critical Thinking	0.308	0.714	-1.124	1.74	0.059	0.668

Table 13 revealed the results on the effect of utilizing AI on the writing proficiency of Grade 11 Senior High School students for grammar, mechanics, organization, vocabulary, and critical thinking, the Beta coefficients represent the estimated effect sizes of utilizing AI on each aspect of writing proficiency. None of the p-values for any aspect of writing proficiency are below the alpha level of 0.05. This means that there is no statistically significant effect of utilizing AI on the writing proficiency of Grade 11 Senior High School students across these specific aspects

However, the result of this study which showed that there is no significant effect of AI writing assistant to the writing proficiency of senior high school students can be explained by the Generalization of Experience theory which stated that transfer of learning to similar or related task can be positive, negative, or neutral. In the case of this study, since there is no significant effect found, transfer of learning can be denoted as neutral.

Another factor that could possibly affected the result is the amount of time respondents are exposed to AI writing assistants. The Technology Acceptance Model asserts that one of the factors that may affect technology's effect is its utilization. Since there was time constraint, students were subject to the intervention for only a short period of time.

4. Conclusion and Recommendations

Based on the initial findings, the following conclusions were drawn:

1. There is a significant difference on the results of the Diagnostic Test and Formative Test, as well as the results of the Formative Test and Summative Test on Writing Proficiency. Therefore, the null hypothesis should be rejected.

The significant differences found among the tests conducted revealed that over time, changes happen to the writing proficiency of senior high school students after the use of AI writing assistants. However, there were also differences in the writing proficiency results of the control group who did not receive any intervention of AI writing assistants. Possibly, this is caused by teacher's feedback on written outputs.

Consequently, conclusion may be made that there is no difference between AI-generated feedback and human-generated feedback.

2. There is a significant difference found on the writing proficiency of the experimental groups and control group in terms of grammar usage and critical thinking skills. However, the results also showed that there is no significant difference on the writing proficiency of the experimental groups and control groups in terms of mechanics, organization, and vocabulary. Due to this, the null hypothesis cannot be completely rejected.

This difference in writing proficiency in terms of grammar usage between the experimental and control groups might be caused by the respondents' perceived usefulness of AI writing assistants in grammar correction. Respondents in the experimental group strongly agreed that AI writing assistants helped them catch grammatical errors in their written outputs. This belief may have caused students to have a stronger sense of confidence in the corrections made by AI writing assistants, thus, leading them to pay more attention to feedback on grammar made by these tools. A significant difference in critical thinking was also found, as this skill relates to all aspects of writing, including grammar usage. Using correct grammar and revising errors in writing requires the use of critical thinking skills.

Out of the five indicators of writing proficiency, only two areas were found to have significant differences between the experimental and control groups. Since most writing proficiency areas did not show any significant difference, the researcher concludes that there is no difference in the writing proficiency of the experimental and control groups. This result can explain why even the control group, who did not receive AI writing intervention, had significant differences in the three tests conducted, similar to the result of the experimental groups who received AI writing interventions.

3. The results of the study revealed that there is no significant effect of AI writing assistants to the writing proficiency of Grade 11 senior high school students of Cristobal S. Conducto Memorial Integrated National High School. This means that the null hypothesis should be accepted.

This result may be caused by the time constraint experienced during the conduct of the study, leading to short period of time respondents were exposed to AI writing assistants as intervention in writing. Although there was difference found in terms of grammar and critical thinking, it was not enough to establish the notion that AI writing assistants have significant effect to writing proficiency. Another factor that could possibly affect this is the errors found by previous studies on different AI writing assistants features. Inaccuracies in the feedback generated by these AI writing assistants can also affect possible improvements in the students' writing proficiency.

Upon the presentation of the findings and conclusion, the following recommendations are made.

1. Further Studies on Writing Proficiency Focusing on Grammar Usage and Critical Thinking

With the result of the study revealing significant differences in writing proficiency in terms of grammar usage and critical thinking skills, it would be beneficial for teachers to conduct further studies focusing in this areas. Limiting the extent of future researches into these two areas of writing proficiency can result to a more targeted intervention which can possibly be beneficial to students.

2. Investigation of the Integration of AI Writing Assistants in Writing Activities

Since there is a difference between the experimental and control groups, teachers and policy makers may take into account integrating AI writing assistants in completing writing activities in class. This does not mean completely eliminating feedback from teachers, but utilizing AI writing tools as aid in checking students' work. This may also involve guided and controlled use of AI writing assistants in class, while promoting critical thinking in accepting corrections made by these tools.

3. Conduct of Longitudinal Study Design

Due to time constraint experienced by the researcher, students were only exposed to AI writing assistants as intervention in writing for a short period of time. This may have led to the absence of significant effect found in the study. Future researchers may consider conducting longitudinal study design to observe long-term effects of utilizing AI writing assistants to the writing proficiency of students. Tracking their progress for a considerable amount of time may lead to a more in-depth analysis of the effect of using AI writing assistants.

Reference:

- Ravichandran, S., Kretovics, M. A., Kirby, K., & Ghosh, A. (2017). Strategies to address English language writing challenges faced by international graduate students in the US. Zenodo (CERN European Organization for Nuclear Research). <https://doi.org/10.5281/zenodo.570033>
- 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.
- Moonpreneur. (2023, July 27). AI's Negative and Positive Impact On Education Industry. Moonpreneur. <https://moonpreneur.com/blog/ai-negative-and-positive-impact-on-education-industry/>
- Moonpreneur. (2023, July 27). AI's Negative and Positive Impact On Education Industry. Moonpreneur. <https://moonpreneur.com/blog/ai-negative-and-positive-impact-on-education-industry/>