

Artificial Intelligence and Writing: An Analysis of AI's Impact on Writing and Grammar Ability

Anirudh Prodduturu

Monroe Township High School, USA

ABSTRACT

Artificial Intelligence, commonly abbreviated as AI, is a new technology that can impact the lives of billions of people given its remarkable usefulness in everyday tasks. Specifically, chatbots, which are AI-based tools that can communicate with a human and complete tasks such as creating a list, writing a paragraph, or summarizing a story, are becoming a popular alternative to manually crafting essays or any form of writing among students. Due to this, hypothetically, students may be getting less practice actually writing essays themselves, which could present a problem in the future because students' writing ability (or grammar ability) could be affected. This study aims to study the influence that extensive AI use for writing has on students at a specific high school in New Jersey. A quantitative correlational survey method was conducted to obtain the data necessary for the study, and a total of 102 responses were collected. The study found that there was no significant influence that usage of AI for writing had on students' grammar ability at the specific high school, as evidenced by the linear correlation coefficient value of 0.179. This study could help to inform educators and policymakers on how to deal with AI usage by students for in-school and out-of-school writing.

Introduction

The recent creation of LLMs (Large Language Models) has completely changed the fields of AI (artificial intelligence) and writing. LLMs are a form of GAI (generative AI) that can understand and communicate through general, human-like language. LLMs are well-regarded for being able to quickly generate responses with large amounts of text within several seconds, something humans cannot do nearly as efficiently. OpenAI's ChatGPT, a popular type of LLM, has exploded in usage, causing speculators to question its applicability to everyday tasks, such as writing. As described in a study by Kohnke et al., although ChatGPT has demonstrated its unparalleled capabilities in producing text, critics have noted that it occasionally struggles to utilize proper grammar and provide accurate details (Kohnke et al., 2023). Assuming that people may use LLMs to generate writing with the presumption that information cited and grammar used are accurate, it is imperative to understand LLMs' impact on one's writing and grammar ability. Understanding this is critical for academia and industries that rely on humans' writing ability (Salvagno et al., 2023). A study that will examine this phenomenon is crucial to understanding the benefits and challenges that LLMs pose and how they can be integrated into our daily lives to maximize what is gained, rather than lost.

Literature Review

Grammar and Writing Ability

One crucial concept to grasp is the relationship between writing and grammar ability. A paper by Muhsin, which aimed to explore this correlation through a descriptive analysis study, found that there was a significant correlation between the students' grammar and writing ability. Muhsin's study, which used the Pearson Product Moment Correlation Coefficient to determine this correlation, focused on 23 second-year students. He concluded that mastery of

sentence structure, which he defines as "writing ability", is correlated with students' ability to produce a grammatically correct paragraph. He states, "There was a significant correlation between the students' grammar and writing ability" (Muhsin, 2015, p. 14). This paper is significant as it provides a basis for understanding how writing ability can be measured, paving the way for a quantitative study that can analyze one's grammar proficiency instead of writing proficiency through a qualitative method.

In comparison, a study by Fahuzul et al. examined the correlation between writing ability and grammar fluency for 8th-grade students. The students were asked to complete a straightforward multiple-choice grammar mastery test and a writing assignment where they were tasked with composing three paragraphs (Fahuzul et al., 2022). The researchers found a very slight to no correlation between 8th graders' grammar level and their ability to write several paragraphs, which is a result that is not entirely consistent with Muhsin's conclusion. The discrepancy between the two studies may be due to the age range that was studied in the papers. In Muhsin's study, the writers were older and, therefore, more developed, so good writers for their age may be more familiar with the command of the English language. In contrast, the 8th graders that were examined in Fahuzul et al.'s study may have needed to be more developed in terms of grammar ability and slightly less representative of the students that will be researched in this study.

Both studies, despite their differing conclusions, agree on the importance of grammar fluency for success in the real world (Fahuzul et al. 2022; Muhsin, 2015). This agreement reinforces the idea that understanding the impact of LLMs on writing/grammar ability is crucial, thereby justifying the main goal of this study.

Potential of LLM Use for Writing

Some papers offer insight on how LLMs can be used to develop the writer. One such study by Fitria examined how well ChatGPT can write English essays. She employed a descriptive and qualitative method to analyze ChatGPT's use of tenses, voices, and sentence structure. Fitria stated, "This chatbot can also help [writers] understand a topic better because it can explain things in an easy-to-understand way" (Fitria, 2023, p. 55). Based on this conclusion, it is evident that writers can use ChatGPT as a tool to get information regarding a topic. This study is important because it demonstrates how ChatGPT can aid writers in their writing, meaning it does have the potential to improve writing.

Another paper by Hwang et al. researched the potential of AI in education and writing courses. They found that it has the potential to revolutionize education because it can provide ideas, engage students, and promote creativity inside classrooms (Hwang et al., 2023). The researchers concluded that LLMs have the potential to improve the writing process and overall education of many students if used correctly. This research can also be applied to writing that is not done in classrooms since LLMs will offer the same benefits if used in the same way. Fitria's study relates to Hwang et al.'s because they agree that ChatGPT is a tool that can improve writing under certain circumstances. They thus offer a positive outlook on how AI can be used as an aid and, therefore, improve the quality of writing that is produced.

Downsides of LLM Use for Writing

While some papers discuss how LLMs can be used to improve writing, others argue that the quality of education and writing will decrease if LLMs are used. A paper by Dergaa et al. discussed LLMs' implications on academic writing. They specify that using ChatGPT for writing can actually harm the writer if used irresponsibly (Dergaa et al., 2023). Some examples of using LLMs irresponsibly are directly copying content generated from the LLM and plagiarizing from sources that are given. They conclude by cautioning against misuse of ChatGPT, emphasizing the importance of transparency, and stating what one used LLMs for in their writing.

Similarly, Kohnke et al. highlight the detrimental effects ChatGPT use can have on writing. They point out ChatGPT's tendency to provide false information (Kohnke et al., 2023). Specifically, they provide an example of ChatGPT giving false information about a language being learned, which is harmful because it can potentially misinform learners who do not verify the information somewhere else. They conclude by warning users of ChatGPT's mistakes and encouraging them to verify all information. Dergaa et al. and Kohnke et al. both state that ChatGPT can

have its flaws. They draw attention to the fact that writers cannot use LLMs irresponsibly, or it could detract from the quality of their writing. They also concede that LLMs have their benefits since they make composing pieces more straightforward, but they reiterate their warnings to the reader.

While all of these pieces analyze the potential/drawbacks of LLM use or the correlation between grammar and writing, there needs to be more academic studies that combine the two concepts into one study. Since LLMs are so new, there has yet to be a defining study that studies the effects of LLMs on writing or grammar ability. Specifically, there needs to be more literature that studies its effects on students. Through the use of a survey, a correlational study that will answer the question, "How does extensive use of artificial intelligence for out-of-school writing affect students' grammar ability at Monroe Township High School?" can close the gap in the body of knowledge.

Methods

In order to find a relationship between a student at Monroe Township High School's (MTHS) AI usage for out-of-school writing and their grammar ability, this study used a survey. The survey asked students to complete 2 tasks:

- Estimate one's AI usage for out-of-school writing
 - o Students answered this question in increments of 20%, meaning they could respond with <20%, <40%, etc.
- Complete an SAT-style grammar test
 - o Students were given 10 sentences and asked to either fill in the blank or replace a word/phrase with the appropriate answer. The students' scores on the grammar test were saved. The questions were modeled after past SAT grammar questions, which reinforces the credibility and accuracy of the evaluation system.

Students randomly took the survey. It was distributed to all students through the online platform Schoology. On Schoology, a Google Forms link to the survey was posted as an update that all students could view. If they wanted to complete the survey, they could select the link and finish it. After the students had completed the survey, the responses were anonymously saved for data analysis on Google Sheets. Initially, the plan was to keep the first question about AI usage, but instead of doing a grammar test, include a free response task where the respondent would answer a prompt. After they answered the prompt, the would have researcher evaluated each participant's writing ability based on their response. However, Muhsin's study that found that writing and grammar ability were correlated with each other meant that utilizing an objective grammar test would eliminate any biases or subjectivity on the researcher's side while measuring similar skills (Muhsin, 2015).

The data collected were all quantitative, meaning they were numerical values. The first variable (students' AI usage for out-of-school writing) was a percentage, and the second (students' grammar test score) was an integer. The values for these variables were organized by respondent and imported into Google Sheets to analyze the data.

Once the data was imported into Google Sheets, a line graph was created with each data point representing every participant in the survey. The independent variable was the student's AI usage for out-of-school writing, which was represented by the x-axis. The dependent variable was the student's score on the grammar test, and was represented by the y-axis. After the data was plotted on the graph, a Google Sheets setting was used to generate a trend line for the graph, and different equations for the trendline were tested in order to find the one with the highest R2 value, which represents how well the trend line fits the data points. Using the trendline and R2 value, the correlation can be determined in the following ways based on the type of trendline equation:

- Linear: If m is positive, the correlation between the independent and dependent variables is positive, and vice versa.
- Quadratic: The a and b values. Depending on the exact values of a and b , the correlation between the variables can be calculated.

This study follows the ethical responsibilities of researchers since the students' responses to the survey were anonymous and confidential. The respondents' personal information was not collected, so no one could see the identity

of anyone that responded. The only information the creator of the survey could access were the anonymous responses and values of each variable. The responses were not shared with anyone, so only the researcher knew the data to maintain a degree of confidentiality. Another ethical consideration was informed consent. The survey participants were told exactly what was expected of them when the link was distributed through Schoology and were informed at the header of the Google Form. AI usage for out-of-school writing was studied because using AI for school assignments is considered cheating, and having students admit to doing so would be considered self-incrimination. Some examples of out-of-school writing are letters, emails, etc.

A correlational study was conducted because the main objective was to determine if there is a relationship between AI use for writing and grammar ability. A correlational study is best for identifying patterns, as Leedy and Ormrod stated in their book *Practical Research: Planning and design* (Leedy & Ormrod, 2019). An experimental study was not chosen because the scope of this study does not allow GAI usage's effect on students' grammar ability to be observed over time.

A survey that tests students' grammar was an effective way to measure their grammar ability because a study by Ingley and Pack used a similar technique to test one's writing (Ingley & Pack, 2023). The researchers used a questionnaire to evaluate one's ability to write, and the second part of this survey was based upon their survey. The grammar test was designed to resemble the SAT grammar test as the SAT is a standardized, reputable test for measuring one's grammar fluency. Questions were taken from past SAT grammar tests, but certain words were tweaked slightly.

One limitation of this study was in relation to the students that were surveyed. Since anyone at MTHS could participate in the survey, the responses may not have been fully representative of the MTHS student population. For example, if a disproportionately large number of seniors took the survey, the grammar scores could be expected to be higher since their grammar ability is, on average, more developed than the grammar of a freshman. To combat this, a diverse array of students must be surveyed to increase the sample size and account for these differences. However, there will still be some sort of variation in the values since it is impossible to survey every student at MTHS, in which case the results would fully represent the sample population.

A delimitation of this study was the population that was surveyed. Students were included in the survey since examining how students' grammar ability is impacted is vital because they will need to potentially use these skills in college and their future lives. For example, having a weak writing foundation at a young age can make it hard for one to effectively complete the numerous writing tasks required in the future, and can restrict them from many opportunities. Adults were not surveyed, as not all adults may frequently complete out-of-school writing tasks, especially relative to students, since students need to write for college essays, program applications, etc. Additionally, students at MTHS, in particular, were selected since if the scope of the study was increased to include all of the students in New Jersey, it would be extremely infeasible to collect data as the resources necessary for that project were not available.

Results and Analysis

The goal of this study was to determine if there was a relationship between GAI (Generative artificial intelligence, as defined previously) usage for out-of-school writing and grammar ability for students at MTHS. All survey responses were anonymous, meaning school emails were not collected along with the survey submission. This was done to encourage students to be honest in their responses, since no one will know who gave each response. After data collection, the survey yielded 102 individual responses out of a population size of approximately 2700 students. The number of respondents per each GAI usage category (as answered in the survey) is shown below.

Table 1. Number of respondents per each GAI usage for out-of-school writing category.

GAI usage as a % of out-of-school writing	<20%	20%+	40%+	60%+	80%+

Number of responses	74	10	8	7	3
----------------------------	----	----	---	---	---

The number of responses where students attested to using GAI for 80% or more of their out-of-school writing was disproportionately less than the number of students using less than 80%. Specifically, only 3 out of 102 total respondents used more than 80%, while 99 out of 102 do not, indicating that very few students use GAI for nearly all of their out-of-school writing. The number of respondents per each grammar test score is shown below.

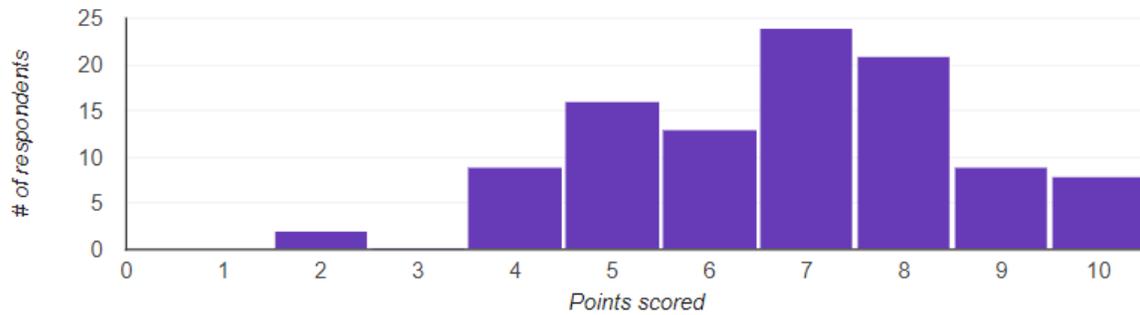


Figure 1. Points distribution for respondents to the grammar quiz.

The grammar test found that a majority of students scored with an accuracy of greater than 7/10. The mean score was 6.81/10 and the median was 7/10, while the range was from 2/10 to 10/10. As mentioned previously, the grammar questions were based on past SAT problems, so it is a proven method for measuring one’s grammar fluency. For each category of GAI usage (<20%, 20%+, 40%+, etc.), the mean grammar test score was calculated by averaging the test scores of all respondents that had a specific GAI usage for out-of-school writing. This means that the category of students that responded with 80%+ only had three corresponding grammar scores, so those three grammar scores were used to calculate the mean for the category.

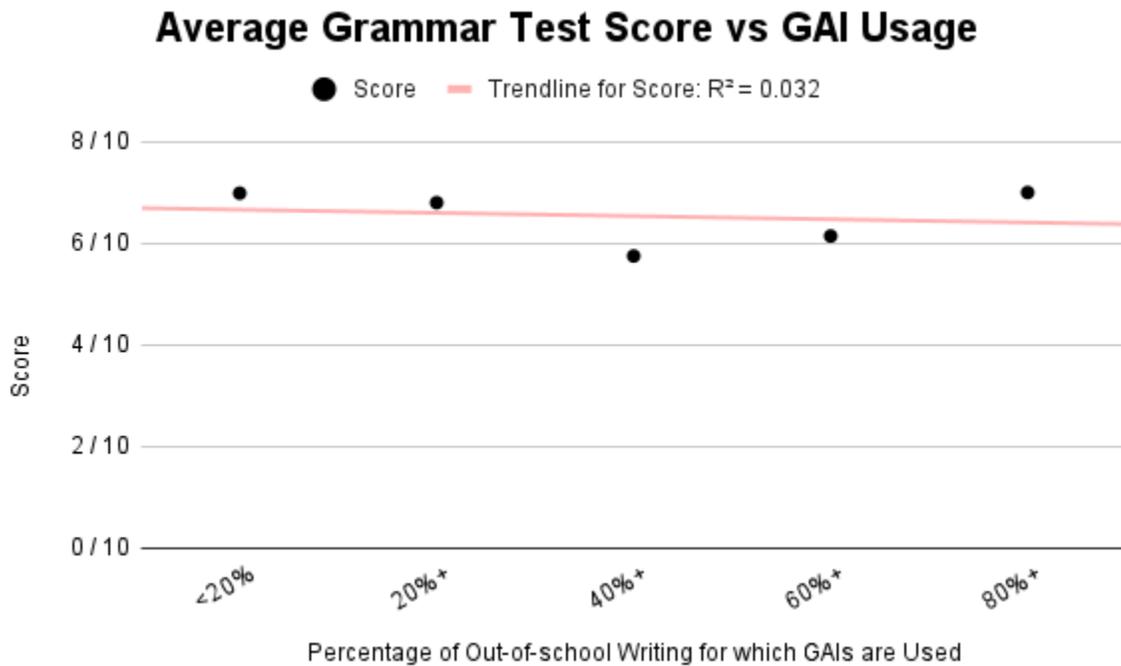


Figure 2. Average grammar quiz score vs GAI usage category (linear).

As shown in the graph, the linear trendline has a slightly negative slope. This suggests that there is not any significant correlation between grammar ability and GAI usage for out-of-school writing. The R2 value, or coefficient of determination, is 0.032. The square root of R2, r , is the Pearson product-moment correlation coefficient, which measures the association between the data points and the trendline (Leedy & Ormrod, 2019). A r value of zero indicates no linear relationship, while a value of one indicates a perfect linear relationship (Puth et al., 2014). Because $r = 0.179$, there is little to no linear relationship between the trendline and dataset. In terms of the research goal, this linear trendline shows that there is essentially no correlation between the two variables, although the trendline has been proven to not represent the data well. Modifying the trendline such that it is modeled by a quadratic function leads to the following graph:

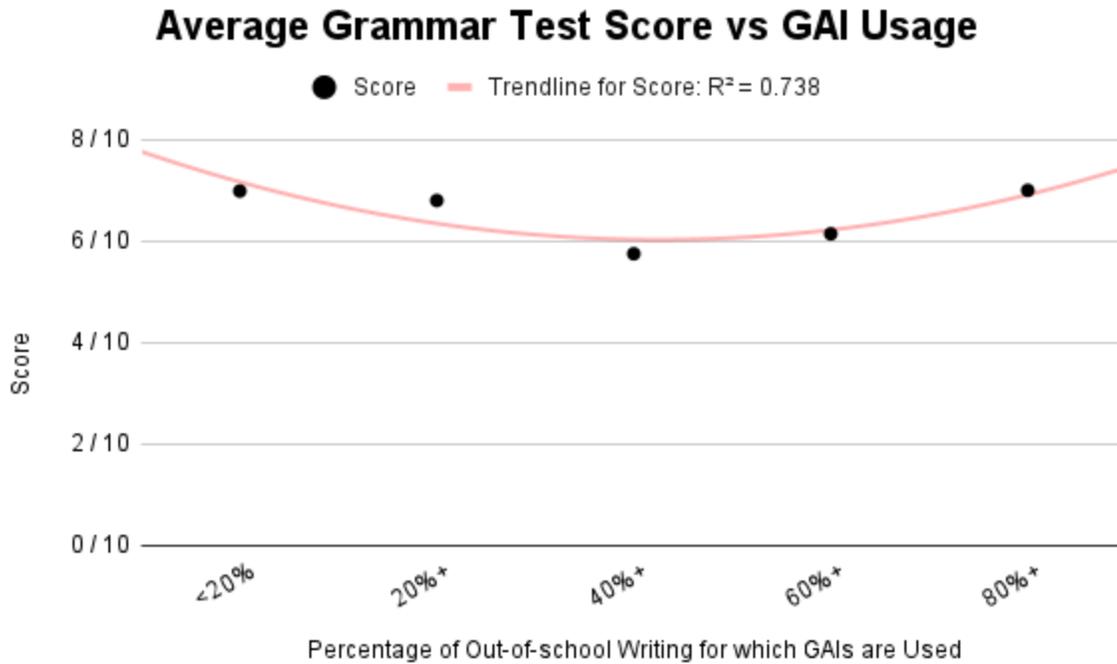


Figure 3. Average grammar quiz score vs GAI usage category (quadratic).

For the quadratic trendline, the coefficient of determination was $R^2 = 0.738$, meaning it represents the dataset better than the linear trendline. The trendline is initially high, bottoms out at 40%+ GAI usage, and then rises again until 80%+ GAI usage. This answers the research question, as it suggests a strong positive correlation between GAI usage for out-of-school writing and grammar ability at the extremes (when it is barely used or almost always used). Conversely, it illustrates a negative correlation between GAI usage and grammar fluency when used for about half of out-of-school writing tasks.

Discussion

In order to comply with ethical guidelines, the survey only measured GAI usage for out-of-school writing. This is because GAI usage for school assignments constitutes cheating, and having students respond to a survey that asked how much they cheat is an extreme ethical violation because it self-incriminates respondents. All respondents to the survey were informed that it was completely voluntary and anonymous. A summary of the study was given prior to students taking the survey, and were given an estimated completion time so that they know whether they would like to do it or not. Additionally, the data was kept confidential and saved in a password-protected school email account. During analysis of the results, no survey responses were excluded from the data, and all data shown in the figures and tables represent the raw data that was collected.

In Fitria's study, she concluded that ChatGPT can aid authors in creating higher quality pieces of writing (Fitria, 2023). This can be evidenced by the fact that students performed well (relative to the median) on the grammar test when they used GAI for 80%+ of out-of-school writing. Extensive usage of GAI in that case may have helped develop the student's grammar and writing skills, leading to a high test score. The quadratic trendline affirms this, since it has a slope that becomes positive as GAI is used for a larger portion of out-of-school writing.

On the contrary, a paper by Dergaa et al. specifies that writers are inclined to use GAI irresponsibly, and irresponsible usage of it can negatively affect their writing ability (Dergaa et al., 2023). This is affirmed by the quadratic trendline. Using the trendline to model grammar test scores if GAI was used for 0% of out-of-school writing, the

score would be higher. Additionally, as GAI usage increases until 60%+, the grammar ability of students continues to decrease, showcasing an inverse relationship between the variables.

The results did not affirm my hypothesis that increased GAI usage will be correlated with a lower grammar score. The research question aimed to determine whether there was a correlation between GAI usage for out-of-school writing and grammar ability, and the trendline confirms that there is no significant correlation between the two. The correlation coefficient for the linear trendline was close to 0, indicating an extremely low correlation between the variables. This means that GAI usage may not have much of an effect on how well students at MTHS can write or follow standard English conventions. The results are significant because they can be applied to the way that students learn. Since there is essentially no correlation, it can be reasonably assumed that extra GAI usage for writing tasks will not negatively affect one's grammar fluency, if used correctly. Although GAI use for writing tasks should not be completely encouraged, the school district can use these results to create policy on how AI can be used for school assignments.

An important limitation in this study was the extent to which the data can be generalized to the entire population of MTHS. There were only three survey responses for the category of 80%+ GAI usage. This means that the datapoint represented in Figure 2 and Figure 3 for 80%+ was just the responses of three people. This is a limitation since it does not effectively generalize the results for students with that specific category. If there were a greater sample size for that category, it would be more generalizable, but that was not the case.

Conclusion

Findings

As found previously in the discussion and analysis section of this study, GAI (Generative AI) usage for out-of-school writing had no major correlation with students' scores on the grammar test. From this, it can be reasonably inferred that GAI usage does not significantly affect Monroe Township High School (MTHS) students' writing and grammar ability. This does not affirm the contrasting viewpoints of Fitria, who predicted that AI usage would correlate with an increase in authors' quality of writing, or Dergaa et al., who suggested that authors using GAI would use it irresponsibly and therefore negatively affect their writing ability. (Fitria, 2023; Dergaa et al., 2023). In certain scenarios these may have been the cases, but no consistent and conclusive evidence justifies one interpretation over the other. The trends deduced from the survey may be applied to the entire MTHS student population because the responses to it were from a random sample of students, and steps to mitigate sampling bias were taken.

Change in Hypothesis

The initial hypothesis of this study was that increased GAI usage for out-of-school writing would negatively correlate with scores on the grammar quiz. The hypothesis was based upon the interpretation of Dergaa et al., as it was assumed that GAI would be used to simply write for students instead of helping them develop their writing through suggestions and feedback (Dergaa et al., 2023). As mentioned, this study found that there was no significant correlation between the 2 variables, meaning that the hypothesis was not supported by this study. This outcome could be because students may not be irresponsibly using GAI for out-of-school writing as much as initially predicted, meaning that their grammar skills won't be negatively affected as much.

Limitations

An important limitation in this study stems from the fact that only three respondents attested to using GAI for more than 80% of out-of-school writing. This is in stark contrast to the 74 respondents who used GAI for less than 20% of

out-of-school writing. The significantly lower number of respondents for that category can be attributed to a lower number of students using GAI for almost all of their out-of-school writing. In anticipation of a gradually lower number of respondents per GAI usage category as it increases, a large sample size was used (102 responses). However, the number of responses for the 80%+ category still yielded only 3. This can skew the results of the survey since the sample size of the 80%+ category was low enough that its mean grammar score may not accurately represent the MTHS student population, which is essential to determining the extent to which GAI usage for out-of-school writing and grammar ability is correlated.

Further Research and Implications

The results generated from this study can be applied to educational settings. Although this study only measured GAI usage for out-of-school writing, meaning school assignments were not a factor, both academic and out-of-school writing constitute a form of writing and can therefore be extrapolated to each other. As described in a study by Hwang et al., action needs to be taken by educational institutions to set boundaries on AI usage (Hwang et al., 2023). This study can help educators at MTHS decide on ways to moderate GAI usage for school writing assignments, given that there is no particular trend between increased GAI usage for out-of-school writing and grammar ability. Furthermore, additional research can be conducted to determine GAI usage's effect on other aspects of English. Another study upon the same population can be used to find the correlation between GAI usage and reading comprehension, for example. This knowledge can further allow educators to understand the implications of GAI usage and how it should be regulated. This study contributes to the somewhat limited body of knowledge on the impact of AI usage on students' learning.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

- Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615-622. <https://doi.org/10.5114/biolsport.2023.125623>
- Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching*, 12(1), 44-58. <https://doi.org/10.15294/elt.v12i1.64069>
- Hwang, G.-J., & Chen, N.-S. (2023). Editorial Position Paper: Exploring the Potential of Generative Artificial Intelligence in Education: Applications, Challenges, and Future Research Directions. *Educational Technology & Society*, 26(2). <https://www.jstor.org/stable/48720991>
- Ingle, S. J., & Pack, A. (2023). Leveraging AI tools to develop the writer rather than the writing. *Trends in Ecology & Evolution*, 38(9), 785-787. <https://doi.org/10.1016/j.tree.2023.05.007>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>

Leedy, P. D., Ormrod, J. E., & Johnson, L. R. (2019). *Practical research: Planning and design* (12th ed.). Pearson.

Muhammad Fahuzul, Badriyah Ulfah, & Desi Surayatika. (2022). The Correlation Between Grammar Mastery and Writing Ability of The Eighth Grade Students. *ULIL ALBAB : Jurnal Ilmiah Multidisiplin*, 1(12), 4394–4403. <https://doi.org/10.56799/jim.v1i12.1098>

Muhsin, M. A. (2015). The correlation between students' grammar knowledge and writing ability.

[https://www.researchgate.net/profile/Muh-Arief-](https://www.researchgate.net/profile/Muh-Arief-Muhsin/publication/303761855_THE_CORRELATION_BETWEEN_STUDENTS'_GRAMMAR_KNOWLEDGE_AND_WRITING_ABILITY/links/57512ae708ae1c34b39caa80/THE-CORRELATION-BETWEEN-STUDENTS-GRAMMAR-KNOWLEDGE-AND-WRITING-ABILITY.pdf)

[Muhsin/publication/303761855_THE_CORRELATION_BETWEEN_STUDENTS'_GRAMMAR_KNOWLEDGE_AND_WRITING_ABILITY/links/57512ae708ae1c34b39caa80/THE-CORRELATION-BETWEEN-STUDENTS-GRAMMAR-KNOWLEDGE-AND-WRITING-ABILITY.pdf](https://www.researchgate.net/profile/Muh-Arief-Muhsin/publication/303761855_THE_CORRELATION_BETWEEN_STUDENTS'_GRAMMAR_KNOWLEDGE_AND_WRITING_ABILITY/links/57512ae708ae1c34b39caa80/THE-CORRELATION-BETWEEN-STUDENTS-GRAMMAR-KNOWLEDGE-AND-WRITING-ABILITY.pdf)

Puth, M., Neuhäuser, M., & Ruxton, G. D. (2014). Effective use of Pearson's product–moment correlation coefficient. *Animal Behaviour*, 93, 183–189. <https://doi.org/10.1016/j.anbehav.2014.05.003>

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care (London, England)*, 27(1), 75. <https://doi.org/10.1186/s13054-023-04380-2>