



## THE IMPACT OF USING ARTIFICIAL INTELLIGENCE (DEEPL) ON ELEVENTH GRADE STUDENTS' WRITING SKILLS AT MAN KARANGASEM

Wahida Syfa Maryam,<sup>1</sup> Made Hery Santosa<sup>2</sup>, Kadek Sintya Dewi<sup>3</sup>

*wahida@undiksha.ac.id*<sup>1</sup>  
*mhsantosa@undiksha.ac.id*<sup>2</sup>  
*sintyadewi@undiksha.ac.id*<sup>3</sup>

**Universitas Pendidikan Ganesha, Bali**

### ABSTRACT

The purpose of this study is to determine whether the use of DeepL in learning can have an impact on students' ability to write English and to find out how the implementation of DeepL impacts students' writing ability. This study uses a mixed method in helping researchers to obtain data, namely Quasi Experimental Non-Equivalent (Pretest and Posttest) Control Group Design on quantitative pre-test and post-test student provision as a form of quantitative data and interviews for qualitative data. The sample involved in this study was 56 students who were divided into two groups: the experimental group, which was the group that received the treatment and also the control group where the researcher conducted conventional learning activities. The use of deep learning in English learning does not provide sufficient impact even though in the acquisition of interview data students have an interest in DeepL. As a result of this research, the researcher assumes that the use of AI-based applications in learning still needs more development and attention.

***Keywords: Artificial Intelligence, Deep, Writing, Writing skills, Conventional***

### INTRODUCTION

Artificial Intelligence as part of Information Technology is present as a solution in education today. Artificial Intelligence is an application related to computer



programming to do something that in a human point of view or can be understood how to make computers able to do things that can currently be done better than humans. According to (Prastiwi & Pujiawati, 2019) Artificial Intelligence (AI) is a science of computer that focuses on the ability of machines to imitate human behavioral intelligence. AI allows computers to process a lot of information and data and provide computer-based conclusions in a relatively short and fast time. The scope and implementation of Artificial Intelligence is very broad, including in the field of education. One of the main ways AI will impact education is through the application of a greater level of individualized learning. Some of this is already happening through a growing number of adaptive learning programs, games, and software. The system responds to student needs, places greater emphasis on specific topics, repeats things that students have not yet mastered, and generally helps students to work at their own pace, whatever it is you. Students can get additional support from AI tutors. While there are obviously things that human tutors can't offer that machines cannot, at least not yet, the future can see more students being mentored by tutors who only exist in zero and one. Several artificial intelligence-based tutoring programs already exist and can help students.

Within the education sector, there has been increased application of artificial intelligence, going over and above the conventional understanding of AI as a supercomputer to include embedded computer systems (Chen et al., 2020). AI can change the role of teachers. There will always be a role of the teacher in education, but what that role is and what it contains may change due to new technologies in the form of intelligent computing systems. As discussed earlier, AI can take over tasks such as assessments, can help students improve learning, and can even be a substitute for tutoring in the real world. However, AI can be adapted to many other aspects of teaching as well. AI systems can be programmed to provide expertise, serve as a place for students to ask questions and find information or even potentially replace teachers for very basic subject matter. However, in many cases, AI will shift the role of the teacher to facilitator.

In this day and age, digital tools, such as online dictionaries, spelling and grammar checkers, and search engines are ubiquitous and can aid the process of writing. In this study, researchers used the DeepL application as a research object material that attracted the attention of researchers. The application is an application that still has few users because it is not as famous as other translation applications such as Google Translate, U-Dictionary, and others. The application attracted the attention of researchers because there is a new feature that other foreign language translation applications do not have.



The ability to write is one of the demands in learning English. However, writing is also a skill that is quite difficult for high school students to deal with. For this reason, there needs to be a learning model that is in accordance with the times and the problems that students are facing. The use of mobile devices as one of the learning media has begun since the development of E-Learning (Electronic Learning) as a learning step to be able to facilitate students so that learning can be done anywhere and anytime. Mobile-based learning makes it very easy for students to come up with learning situations where they present a flexible, effective, interactive, and interesting learning so that this learning is more meaningful. As is known, English is a second language for Indonesian students. This is the background to implement a special learning model so that students can capture learning well because this is also related to students' motivation and attitudes in learning a language that is not their language. It is a fact that students still have difficulty in writing. There are several factors that affect students' writing skills; First, students do not master the language components well such as; vocabulary, punctuation, phonology and grammar. Second, students do not want to practice their writing skills. (Wahyuni & Etfita, 2018)

In this case, the researcher wants to conduct research using an experimental method that aims to experiment whether the use of DeepL has an impact on students' writing skills. By applying several theories from experts and also the support from the condition of students at MAN Karangasem who also experience problems in learning to write, as well as the use of AI as a learning medium in schools, encourage the author to conduct further research. The research question discussed in this study is there any effect of DeepL implementation on grade eleventh students' writing skills.

## **METHOD**

This research used a mixed methods design, in which the researcher collects and analyzes both quantitative and qualitative data. The researcher then integrates the information into an experiment or intervention study (Creswell & Creswell, 2014). This research was undertaken in MAN Karangasem because this school has adopted several AI media based on learning. There were 56 eleventh grade students who were separated into two groups, the experimental group and the control group. These groups were selected using simple random sampling, where the researchers used a lottery to decide the group. Each group consists of 28 students, and they were tested after using DeepL to see the effect of DeepL on their writing skills. To prove that DeepL was effective to be used in the writing learning process, the written test was shared to be tested for two groups. Before conducting the post-test, the treatment was implemented in six sessions.



In analyzing the data gained, descriptive and inferential statistics were utilized to prove that DeepL affected eleventh grade students' writing skills.

## RESULTS

The research used a written essay test for pre-test and post-test to analyze the effectiveness of DeepL on eleventh grade students' writing skills. Before implementing the treatment, the researchers administered the pre-test to the experimental and control groups. After that, the researchers completed six meetings to implement DeepL in the learning process before conducting a post-test for those groups. The result of the post-test showed that the mean scores of the experimental group were higher than those of the control group. The average value of the groups that received treatment using DeepL and those who did not receive treatment had a difference that was not much different. The experimental group had an average score of 81.067. Meanwhile, the control group had a mean score of 80.214. The difference between the means of the two groups is only 0.853. Thus, it can be demonstrated that DeepL has an insignificant effect on the students' writing ability.

## DISCUSSION

In gathering and analyzing the data of the study, the researcher has gone through many procedures. First, the sample of this study was selected by using the criteria or requirements set by the researchers, such as having a smart phone and being able to use it, having permission from the school and parents to be samples of this study, and being willing to participate in this study. Second, the researchers used simple random sampling to determine the experimental and control groups of this study. Finally, after making sure that all these steps were done, the researcher started to provide treatment for these groups. The experimental group was the group that was treated by using DeepL during the learning process. Meanwhile, the control group was the group that used printed book during the learning process. In providing treatment for these groups, the researchers used six sessions to make sure that the students had experience of learning by using the treatment. After completing six meetings, the researchers conducted a post-test to investigate whether DeepL is affecting the student's skills or not. The researchers used SPSS v.26 to analyze the results of the post-test.

### Descriptive Statistics

A statistics program or SPSS v.26 was used to analyze the data descriptively after calculating the post-test scores. It was used to discover the mean, median, mode, standard deviation, variance, range, minimum, and maximum form the pre-test and also



Authors: Maryam, W. Syfa; Santosa, Made Hery; Dewi, Kadek Sintya

post-test scores of the eleventh grade. The post-test scores were employed in this descriptive analysis.

Table 1. Descriptive Analysis Pre-test Score

		Experimental Group Pre-Test Score	Control Group Pre-Test Score
N	Valid	28	28
	Missing	0	0
<b>Mean</b>		78.046	76.586
<b>Median</b>		77.800	78.600
<b>Mode</b>		76.0	79.6
<b>Std. Deviation</b>		5.5651	7.6756
<b>Variance</b>		30.970	58.915
<b>Range</b>		18.2	39.2
<b>Minimum</b>		68.6	50.0
<b>Maximum</b>		86.8	89.2

Table 2. Descriptive Analysis Post-test Score

		Experimental Group Post-Test Score	Control Group Post-Test Score
N	Valid	28	28
	Missing	0	0
<b>Mean</b>		81.067	80.214
<b>Median</b>		80.200	81.000
<b>Mode</b>		78.0	81.0
<b>Std. Deviation</b>		5.7670	8.3807
<b>Variance</b>		33.258	70.237
<b>Range</b>		20.8	29.0
<b>Minimum</b>		68.0	61.0
<b>Maximum</b>		88.8	90.0

The table above showed that the mean of both groups was different, in which the on experimental group has 78.046 mean scores for pre-test and 81.067 for post-test. Meanwhile, control group has 76.586 mean scores for pre-test and 80.214. In addition, the experimental group has a lower median score than the control group. The experimental group has a median score of 77.800 for pre-test score and 80.200 for post-test score, while the control group has a median score of 78.600 for pre-test score and 81.000 for post-test. However, researchers need to calculate the overall data to see the significant difference between the two groups.



To see the variation of the whole data around the median and mean, standard deviation among those groups were different. The standard deviation of the experimental group was 5.5651 for pre-test scores and 5.7670 for post-test, while the control group was 7.6756 for pre-test and 8.3807 for post-test. The scores obtained demonstrate the use of DeepL in students' writing learning only provides a slight change. The difference can be seen from the score obtained by the experimental group is slightly higher than the control group.

### Inferential Analysis

After analyzing the data through descriptive statistics, the inferential statistics then analyzed to achieve the final result of this study. In this case, independent sample Mann-Whitney U-test used to compare the two groups. To conduct the U -test, a requirement had to be done by the researchers, namely normality and homogeneity test among both groups. Specifically, the researchers utilized Shapiro-Wilk to test the normality among the groups, and utilized Levene Statistic to test the homogeneity. The normality test was categorized as distributed if the significance value (.Sig) was above the value of .05. Meanwhile, the homogeneity is categorized as a data can give a normal distribution if the significance value was not lower than .05

#### a) Normality Test

To test the normality among the group, Shapiro-Wilk used because the sample of this study was less than 50 participants (Pallant, 2017), in which the researchers only used 28 students in each group.

Table 3. Normality Test Post-test Scores

	Kolmogorov-Smirnov			Saphiro-Wilk			
	Name of Group	Statistic	df	Sig.	Statistic	df	Sig.
Post-test Score	Experimental Group	.162	28	.058	.904	28	.015
	Control Group	.168	28	.041	.908	28	.017

Based on the results of the calculations that have been carried out, the table shows that the student post-test data shows a value below .05 (.015), this indicates that the data distribution is not normal.

#### b) Homogeneity Test

The homogeneity test was conducted through Levene Statistics to see whether the samples were homogeneous or not. To test homogeneity among the groups, the researcher used student's scores to test it.



Authors: Maryam, W. Syfa; Santosa, Made Hery; Dewi, Kadek Sintya

Table 3. Homogeneity Test Post-test Scores

		Levene Statistic	df1	df2	Sig.
Post-test Score	Based on Mean	5.068	1	54	.028
	Based on Median	4.105	1	54	.048

From the table it can be seen that the homogeneity test value obtained is below .05 (.028). This indicates that the data is not homogeneous. These results are the same as the previous normality test results.

#### c) Independent Sample Mann-Whitney U-test

After knowing that both data are not normal and not homogeneous, the next calculation process uses Mann-Whitney because the data is not parametric. The independent sample Mann-Whitney U-test is used to calculate whether there is a significant impact on the treatment given but in this case for non-parametric data.

Table 4. Independent Sample Mann-Whitney U-test

Total N	56
Mann-Whitney U	383.500
Wilcoxon W	789.500
Test Statistic	383.500
Standard Error	60.995
Standardized Test Statistic	-.139
Asymptotic Sig. (2-sided test)	.889

The table above revealed that the value of Asymp Sig. 2-sided test was .889 (higher than .05). the value indicates there is no effect in using DeepL on eleventh grade students' writing learning.

Based on the findings and discussion presented above, this study shows that the use of DeepL as a learning media does not affect students' English language skills, especially in writing skills. This can be influenced by the lack of familiarity of DeepL in a society that is mostly more accustomed to GT. This is supported by the familiarity theory initiated by Komiak & Benbasat (2006) which has revealed that Familiarity increases the intention to adopt through cognitive trust and emotional trust. However, the use of AI can certainly still be applied in schools because considering the use of technology in learning in the current era must be applied frequently but must still pay attention to the work system of the MT (Derian Ramdhani, 2021). This is related to



statement conducted by Radhaswati and Santosa (2022) that Integrating innovative media and technology needs to be implemented in the school. Muarif et al., n.d. also stated that the use of AI in learning can affect students' motivation and interest in learning. The other factor is the online learning also makes them easier to use language tools (Yuliasiti et al., 2023). With advanced technology, students can access learning materials in a more interactive and interesting way. On the other hand, this must be aligned with the development of technology by educators. artificial intelligence (AI) creates opportunities and challenges for teachers and lecturers of the English language of the future (Subiyantoro et al., 2023). Because at the core, English teachers must be equipped with learning materials, strategies and methods that support student development (Hery Santosa et al., 2020).

## CONCLUSION

DeepL Write is a Web-application that can be used to help fix English sentences by fixing the user's grammar and providing alternative vocabulary directly. This application can be used for all groups, both students and the general public. Based on the results of research that has been carried out, the use of DeepL in the process of learning to write English students at MAN 1 Karangasem does not have an impact on students' writing skills. This can be seen from the results of non-parametric data calculations. According to descriptive analysis, the mean score of experimental group was higher than the control group. However, when calculating data to calculate normality and homogeneity tests, the data results are not normal and not homogeneous, which means the data is not parametric so researchers use non-parametric methods. With the results shown when calculating normality and homogeneity, it shows that the use of DeepL in the learning process does not have the expected impact. So that the researcher's hypothesis that has been compiled previously which assumes that the use of DeepL is able to have an impact on students' English writing skills at MAN 1 Karangasem is rejected.

The use of DeepL in learning English, especially in learning to write, does not provide sufficient impact even though in the acquisition of interview data students have an interest in DeepL. This is due to the free features that are easily accessible for students. The use of AI-based applications in learning still needs more development and attention. In its application, it also does not release the attention of the teacher as a teacher.



## REFERENCES

- Adhabi, E. A. R., & Anozie, C. B. L. (2017). Literature Review for the Type of Interview in Qualitative Research. *International Journal of Education*, 9(3), 86. <https://doi.org/10.5296/ije.v9i3.11483>
- Cancino, M., & Panes, J. (2021). The impact of Google Translate on L2 writing quality measures: Evidence from Chilean EFL high school learners. *System*, 98, 102464. <https://doi.org/10.1016/j.system.2021.102464>
- Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. <https://doi.org/10.1016/j.system.2020.102408>
- Creswell, J. W., & Creswell, J. D. (2014). *creswell*.
- Derian Ramdhani, A. (2021). *Penggunaan Google Translate dalam Menunjang Pembelajaran Bahasa Inggris Siswa*. <https://www.researchgate.net/publication/351711929>
- Dewi, A. K., Ratminingsih, N. M., & Santosa, M. H. (2020). Mobile-Assisted Task-Based Language Learning, Writing Competency, And Motivation. *JPI (Jurnal Pendidikan Indonesia)*, 9(1), 119. <https://doi.org/10.23887/jpi-undiksha.v9i1.23164>
- Dizon, G., & Gayed, J. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal*, 17(2), 74–92. <https://doi.org/10.29140/jaltcall.v17n2.336>
- Fitria, T. N. (2021). Grammarly as AI-powered English Writing Assistant: Students' Alternative for Writing English. *Metathesis: Journal of English Language, Literature, and Teaching*, 5(1), 65. <https://doi.org/10.31002/metathesis.v5i1.3519>
- Flick Uwe, Kardorff Ernst von, & Steinke Ines. (n.d.). *A Companion to Qualitative Research*.
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
- Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, 4(4), 134–147. <https://doi.org/10.1016/j.ijis.2020.09.001>



- Harmer, J. (n.d.). *Jeremy\_Harmer\_How\_to\_Teach\_Writing\_Longm.*
- Hery Santosa, M., Putu Surya Pratama, I., & Nyoman Adi Jaya Putra, I. (2020). *Developing Android-Based English Vocabulary Learning Materials For Primary School Students.*
- Jacobs, H. L. (1981). Testing ESL composition : a practical approach. *Massachusetts: Newburry House Publisher, Inc, January 1981.*
- Jaya, H., Sabran, S., Idris, M., Djawad, Y. A., Ilham, A., & Ahmar, A. S. (2018). *Kecerdasan Buatan.* Fakultas MIPA Universitas Negeri Makassar.
- Komiak, & Benbasat. (2006). The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents. *MIS Quarterly*, 30(4), 941. <https://doi.org/10.2307/25148760>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175. <https://doi.org/10.1080/09588221.2018.1553186>
- Muarif, J. A., Jihad, F. A., Alfadli, M. I., & Setiabudi, D. I. (2023). 3.+Juhdan. *Anfa Mediatama*, 2. <https://doi.org/https://doi.org/10.572349/seroja.v1i2.548>
- Prastiwi, C. H. W., & Pujiawati, N. (2019). Penggabungan Artificial Intelligence dan Kecerdasan Alami dalam Pembelajaran Keterampilan Menulis Bahasa Inggris. *Prosiding Seminar Nasional Pascasarjana (PROSNAMPAS)*, 2(1), 172–178.
- Radhaswati, I. D. A. A., & Santosa, M. H. (2022). Teachers' Perceptions: the Use of Google Form as a Media to Assess Primary School Students. *EDUTEK : Journal of Education And Technology*, 5(4), 910–924. <https://doi.org/10.29062/edu.v5i4.308>
- Setiyadi, A., Sukirlan, M., & Mahpul, M. (2018). *TEACHING LANGUAGE SKILLS: Preparing Materials and Selecting Techniques.* Graha Ilmu.
- Shivaraju, P. T., Manu, G., Vinaya, M., & Savkar, M. K. (2017). Evaluating the effectiveness of pre- and post-test model of learning in a medical school. *National Journal of Physiology, Pharmacy and Pharmacology*, 7(9), 947–951. <https://doi.org/10.5455/njppp.2017.7.0412802052017>
- Subiyantoro, H., Hartono, R., Fitriati, W., & Faridi, A. (2023). *Prosiding Seminar Nasional Pascasarjana Dampak Kecerdasan Buatan (AI) terhadap Pengajaran Bahasa Inggris di Perguruan tinggi: Tantangan dan Peluang.* <http://pps.unnes.ac.id/pps2/prodi/prosiding-pascasarjana-unnes>
- Van Meter, D. S., & Van Horn, C. E. (1975). The Policy Implementation Process. *Administration & Society*, 6(4), 445–488. <https://doi.org/10.1177/009539977500600404>



Jurnal JOEPALLT

Volume 13 Number 01 Maret 2025

ISSN 2338-3739 (Print)

ISSN 2614-8099 (online)

<https://jurnal.unsur.ac.id/jeopallt>

---

Authors: Maryam, W. Syfa; Santosa, Made Hery; Dewi, Kadek Sintya

Yuliasiti, L. P. S., Santosa, M. H., & Myartawan, I. P. N. W. (2023). An Analysis Of 10th Grade High School Students' Learning Approaches in the Context of EFL Online Learning. *Journal of English Development*, 3(01), 45–53.  
<https://doi.org/10.25217/jed.v1i01.2202>