



The Effectiveness of DeepSeek as a Translation Tool on Students' Reading Comprehension

Ratna Hafti Mustafidatul Ilmiah^{1*}, Yuyun Bahtiar², Iin Baroroh Ma'arif³, Luluk Choirun Nisak Nur⁴

^{1,2,3,4}English Department, Universitas KH. A. Wahab Hasbullah

*Email: haftiilmi@gmail.com

ABSTRACT

This study aims to determine the effectiveness of using DeepSeek AI as a translation tool on students' reading comprehension. This study was motivated by the difficulties experienced by 10th-grade students at MA Unggulan K.H. Abd. Wahab Hasbulloh in understanding English texts. The study used a pre-experimental design with a one-group pre-test–post-test model and involved 35 students selected through purposive sampling. The instrument used was a 20-item multiple-choice test, which had been tested for validity using Aiken's V formula and reliability using Cronbach's Alpha. The results showed that the instrument had very high validity and strong reliability ($\alpha = 0.896$). The normality test showed that the pre-test data were normally distributed, while the post-test data were not, so the analysis of score differences was performed using the Wilcoxon Signed Rank Test. The test results showed a significance value of < 0.001 , indicating a significant effect on students' reading comprehension after using DeepSeek AI. These findings prove that DeepSeek AI can help students understand texts better through accurate and contextual translations, thereby facilitating the process of interpreting and understanding the content of the reading material. Thus, this study concludes that DeepSeek AI is effective as a translation tool for students' reading comprehension. The integration of artificial intelligence technology in English language learning can be an alternative solution to support the strengthening of students' literacy and reading strategies in the digital age.

Keywords: DeepSeek; AI Translation; Reading Comprehension

INTRODUCTION

In the era of globalization, English proficiency has become an important requirement for students to access information, academic knowledge, and international communication. English serves as the global language that dominates the fields of education, technology, trade, and science (Mauliska & D'Angelo, 2024). Good English language skills enable students to access a wider range of learning resources and increase their academic and professional competitiveness. However, the level of English proficiency in Indonesia is still relatively low. The English Proficiency Index (EPI) report by English First (2024) shows that Indonesia ranks 80th out of 116 countries with a score categorized as low. This condition shows a gap between the expected English proficiency and the actual abilities of students in the field.

In English language learning, reading skills play a very important role because they are the main means of obtaining information and academic knowledge. Sahmal et al. (2022) state that reading is an essential skill among the four language skills because most sources of knowledge are presented in the form of English texts. Reading is not just about recognizing words, but also involves the ability to understand the meaning of the text in depth. Giawa (2022) and Rahmawati et al. (2024) emphasize that reading comprehension is related to students' ability to grasp main ideas, interpret information, and draw conclusions from the text they read. Therefore, developing reading comprehension skills is an important focus in English language learning.

However, the reading skills of students in Indonesia are still relatively low. The report of OECD (2023) shows that only about 25% of Indonesian students have reached Level 2 in reading literacy, far below the OECD average of 74%, and almost no students have reached the highest level. This fact indicates a serious problem in reading learning that has the potential to hinder students' academic success. Previous studies have revealed that low reading comprehension is caused by limited vocabulary, low motivation to learn, and difficulty understanding the meaning of texts in context (Prihatini, 2020).

With the development of digital technology, artificial intelligence (AI) has begun to be utilized in education, one of which is through the use of automatic translation tools. AI-based translation tools are considered to help students understand foreign language texts more quickly and efficiently. Google Translate is the tool most widely used by students in the learning process (Bangun & Mustafa, 2021, in Bunga & Katemba, 2024). However, its effectiveness is still debated due to limitations in maintaining context and accuracy of translation (Chandra & Yuyun, 2018). This shows the need for alternative AI-based translation tools that perform better in supporting reading comprehension.

One of the latest innovations in the field of artificial intelligence is DeepSeek AI, an AI-based translation tool that shows competitive performance in text translation. Chen et al. (2025) reported that DeepSeek was able to surpass GPT-4.0 in translation tasks without specific context. However, empirical studies examining the use of DeepSeek AI in an educational context, particularly in improving students' reading comprehension skills, are still very limited. This limitation in research is a gap in previous studies.

Based on preliminary observations at MA Unggulan K.H. Abd. Wahab Hasbulloh, the school has adequate technological facilities, such as computers and internet access. However, the use of AI-based translation tools in reading instruction has not been optimized. Students still have difficulty understanding English texts, which results in low active participation in reading activities. This condition indicates a mismatch between the availability of technology and its pedagogical application.

Therefore, this study aims to examine the effectiveness of using DeepSeek AI as a translation tool on students' reading comprehension skills. This study supports and develops previous research on the use of AI in English language learning, while offering something new by focusing on DeepSeek AI, which has rarely been studied. It is hoped that the results of this study can provide empirical contributions to the development of technology-based reading learning and improve the quality of English language learning.

METHOD

This study used a quantitative approach with an experimental method through a one-group pre-test–post-test pre-experimental design to determine the effectiveness of using DeepSeek AI as a translation tool on students' reading comprehension. The research was conducted at MA Unggulan K.H. Abd. Wahab Hasbulloh, with a population of all 202 tenth-grade students, while the research sample consisted of 35 tenth-grade students selected using purposive sampling based on conducive classroom conditions and diversity in students' English language abilities. The research procedure was carried out in five meetings, beginning with a pre-test to determine the students' initial reading comprehension abilities, followed by three treatments in the form of learning to read recount texts using DeepSeek AI as a tool to translate difficult words or sentences, both in groups and individually, and ending with a post-test to measure reading comprehension abilities after the treatment. The research instrument consisted of a 20-question multiple-choice written test used as a pre-test and post-test with indicators of finding factual information, determining main ideas, understanding vocabulary meanings, identifying references, and making inferences. Data collection techniques were carried out through tests and supported by documentation of learning activities. Data analysis included testing the validity and reliability of the instruments, testing normality using the Shapiro–Wilk, and testing hypotheses using a paired sample t-test if the data were normally distributed or a Wilcoxon Signed Rank test if the data were not normally distributed, with data processing using SPSS version 27 at a significance level of 0.05.

RESULT AND DISCUSSION

This section presents the research findings and a systematic discussion based on the data obtained. The research findings are described first to provide a clear picture of the quality of the instruments and the results of the data analysis. The presentation of results includes validity tests using Aiken's V index, reliability tests using Cronbach's Alpha, normality tests using the Shapiro-Wilk test, and hypothesis testing using the Wilcoxon Signed-Rank Test. Each analysis result is presented in tabular form and followed by a descriptive explanation to facilitate the reader's understanding. Furthermore, in the

discussion section, the research results are interpreted to explain the effect of using DeepSeek as a translation aid on students' reading ability. The discussion also relates the research findings to relevant theories and previous research results, thereby demonstrating the significance and contribution of the study. Thus, this section not only presents data but also provides in-depth analysis to address the research questions accurately and systematically.

Result

This section presents the research results, which include instrument validity testing, reliability testing, data normality testing, and hypothesis testing. The research results are presented in tables and descriptive narratives to facilitate the reader's understanding of the research findings.

- **Validity Test Result**

Validity testing in this study was conducted to determine the level of suitability of the instruments used to measure students' reading abilities. The validity of the instruments was analyzed using Aiken's V index based on the assessments of two validators, namely an English education lecturer and an English teacher at MA Unggulan K.H. Abd. Wahab Hasbulloh. Each validator gave a score to the instrument indicators, and then the scores were converted into Aiken's V values.

Table 1 Criteria Validity

Validation Result	Description
0,77 – 1,00	Very high
0,60 – 0,76	High
0,42 – 0,59	Moderately
0,25 – 0,41	Low
0,00 – 0,24	Very low

The criteria for interpreting instrument validity are presented in Table 1, with a value range of 0.00 –1.00 divided into five categories, namely very low, low, moderate, high, and very high. These criteria are used as a basis for determining whether the instrument is suitable for use or needs improvement.

Table 2 Results of Aiken's V Pre-test Index Analysis

Item	Rater		S1	S2	Amount	n(c-1)	v	Description
	1	2						
1	4	5	3	4	7	8	0,875	Very High
2	5	5	4	4	8	8	1	Very High
3	5	5	4	4	8	8	1	Very High
4	5	4	4	3	7	8	0,875	Very High
5	5	5	4	4	8	8	1	Very High
6	5	5	4	4	8	8	1	Very High
7	4	4	3	3	6	8	0,75	High
8	4	5	3	4	7	8	0,875	Very High
9	5	5	4	4	8	8	1	Very High
10	5	5	4	4	8	8	1	Very High

Table 2 presents the results of the pre-test instrument validity analysis using Aiken's V. As shown in the table, nine items (items 1, 2, 3, 4, 5, 6, 8, 9, and 10) obtained V values ranging from 0,879 to 1,0 which fall into the very high category, while item 7 obtained a value of 0,75, which falls into the high category.

Overall, the Aiken's V coefficients for the pre-test instrument are presented in the table below.

Table 3 Results of Aiken's Pre-test Coefficient Calculation

Item	Rater		S1	S2	Amount	n(c-1)	v	Description
	1	2						
1 -10	47	48	37	38	75	80	0,9375	Very High

Overall, Aiken's V coefficient for the pre-test instrument is 0,9375, as shown in Table 3. This value falls into the very high category, indicating that the pre-test instrument is well designed, relevant, and capable of accurately measuring the research variables, making it suitable for use in data collection.

Table 4 Results of Aiken's Post-test Index Analysis

Item	Rater		S1	S2	Amount	n(c-1)	v	Description
	1	2						
1	5	5	4	4	8	8	1	Very High
2	5	5	4	4	8	8	1	Very High
3	5	5	4	4	8	8	1	Very High
4	4	4	3	3	6	8	0,75	High
5	5	5	4	4	8	8	1	Very High
6	5	5	4	4	8	8	1	Very High
7	4	4	3	3	6	8	0,75	High
8	4	5	3	4	7	8	0,875	Very High
9	4	5	3	4	7	8	0,875	Very High
10	5	5	4	4	8	8	1	Very High

Based on Table 4, the results of the post-test instrument validity analysis using Aiken's V indicate that eight items (items 1, 2, 3, 5, 6, 8, 9, and 10) fall into the very high category with V values ranging from 0,875 to 1,0, while items 4 and 7 obtained a value of 0,75, which falls into the high category. The overall Aiken's V coefficient for the post-test instrument is presented in the table below.

Table 5 Results Aiken's Post-test Coefficient Calculation

Item	Rater		S1	S2	Amount	n(c-1)	v	Description
	1	2						
1 -10	46	48	36	38	74	80	0,925	Very High

Based on Table 5, the overall post-test Aiken's V index is 0.925, which falls into the very high category. indicating that the pre-test instrument is well designed, relevant, and capable of accurately measuring the research variables, making it suitable for use in data collection.

The results of the validity calculations for both instruments indicate that the pre-test and post-test instruments both have a very high category of validity and are suitable for use in collecting data on students' reading comprehension.

- Reliability Test Result

Table 6 Case Processing Summary

		N	%
Cases	Valid	30	100.0
	Excluded ^a	0	.0
	Total	30	100.0

a. Listwise deletion based on all variables in the procedure.

Based on Table 6, the dataset processed in this study consisted of 30 respondents with a validity rate of 100% with no missing data. This indicates that all respondents provided complete answers on every research instrument, both during the pre-test and post-test phases. Thus, the data obtained has a high level of reliability and is capable of representing all respondents, so that the resulting analysis can support the drawing of research conclusions more accurately and objectively.

Table 7 Reliability Statistics

Cronbach's Alpha	N of Items
.896	2

Based on Table 7, the Cronbach's Alpha value of 0,896 falls into the category of high reliability ($\alpha > 0,80$). These results indicate that the test instrument exhibits high consistency and is reliable for data collection.

- Normality Test Result

Table 8 Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Pre-Test	32	91.4%	3	8.6%	35	100.0%
Post-test	32	91.4%	3	8.6%	35	100.0%

Based on Table 8, out of the total 35 students, 32 data points were deemed valid (91,4%) for both the pre-test and post-test, while 3 data points (8,6%) were marked as missing due to the students' failure to take the tests. Consequently, all statistical analyses in this study were conducted using 32 complete data pairs.

Table 9 Test of Normality

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre-Test	.148	32	.074	.954	32	.187
Post-Test	.128	32	.196	.927	32	.033

a. Lilliefors Significance Correction

Based on Table 9, the result of the Shapiro-Wilk test shows that the pre-test data obtained a significance value of 0,187 ($>0,05$) and were therefore determined to be normally distributed. In contrast, the post-test data obtained a significance value 0,033 ($< 0,05$), indicating that the data are not normally distributed. Because one of the data sets did not meet the assumption of normality, hypothesis testing was conducted using the non-parametric Wilcoxon Signed-Rank Test.

- Hypothesis Test Result

Table 10 Hypothesis Test Summary

	Null Hypothesis	Test	Sig.a,b	Decision
1	The median of differences between pre-test and post-test equals 0.	Related-Samples Wilcoxon Signed-Rank Test	<.001	Reject the null hypothesis.
a. The significance level is .050.				
b. Asymptotic significance is displayed.				

Based on Table 10, the results of the Wilcoxon Signed-Rank Test show a significance value of $<0,001$, which is well below the significance level of 0,05. Thus, the null hypothesis is rejected, meaning there is a significant difference between the pre-test and post-test scores. These results indicate that the use of DeepSeek as a translation tool has a significant effect on students' reading comprehension.

Discussion

This discussion aims to determine the research findings regarding the effectiveness of using DeepSeek as a translation tool on the reading comprehension of 10th-grade students at MA Unggulan K.H. Abd. Wahab Hasbulloh. The results of the Wilcoxon Signed-Rank Test showed a significance value of $<0,001$, indicating a highly significant difference between pre-test and post-test scores. Thus, the alternative hypothesis (H1) is accepted, namely that DeepSeek has a significant effect on students' reading comprehension. These findings suggest that DeepSeek can serve as an effective tool in helping students understand English texts.

This effectiveness can be explained by its ability to reduce vocabulary barriers. According to Fadila et al. (2024), Limited vocabulary is one of the main causes of difficulty in understanding reading material. DeepSeek provides quick and contextual translations, allowing students to understand words without having to pause the reading process for too long. This helps students focus more on the text content, such as identifying main ideas and drawing conclusions (King & Stanley, as cited in Nurdiana & Amelia, 2017).

These research results are also consistent with previous studies. Kim & Cha (2023) and Allehyani et al. (2025) found that AI-based translation tools influence improvements in reading comprehension. This study reinforces these findings while offering a new contribution by focusing on DeepSeek, which, according to Weigang & Brom (2025), provides high-quality translation that preserves meaning, thereby helping students understand the text without confusion.

However, there are several limitations, such as a small sample size, a tendency toward high post-test scores, and potential for students to use external aids. Therefore, future research is recommended to use a larger sample size, a design with a control group, and stricter supervision. Overall, DeepSeek has proven effective in helping students improve their reading comprehension. In addition to translating, this tool also helps reduce language difficulties and enhance text comprehension. These results indicate that AI has great potential in education and can be utilized by teachers as a strategy in reading activities.

CONCLUSIONS

Based on the research findings, it can be concluded that the use of DeepSeek as a translation tool significantly contributes to improving the reading comprehension of 10th-grade students at MA Unggulan K.H. Abd. Wahab Hasbulloh. The use of this technology addresses the initial problem identified in the preliminary stage, namely, students' low ability to understand English text. Through fast and contextual translation support, DeepSeek helps students overcome language barriers, making the text comprehension process more effective and efficient.

Research findings indicate an improvement in students' abilities following the use of DeepSeek in learning, suggesting that AI-based technology can serve as a learning medium that supports reading comprehension more effectively. Thus, DeepSeek functions not only as a translation tool but also as a means of facilitating students' cognitive processes in understanding reading material.

However, this study still has limitations, particularly regarding the limited sample size and research design that did not include a control group. Therefore, future research is expected to develop a stronger experimental design, involve a broader range of participants, and explore the use of DeepSeek over the long term as well as in other language skills. With such developments, the use of AI in learning is expected to make a broader contribution to the world of education.

REFERENCES

- Allehyani, B., Almashy, A., Jamshed, M., & Abdul, S. B. (2025). Measuring the Impact of Meta-AI on English Reading Comprehension Score Enhancement : A Study Within Social Media Application. *Theory and Practice in Language Studies*, 15(2), 615–623. <https://doi.org/10.17507/tpls.1502.31>
- Bunga, E. L. M., & Katemba, C. V. (2024). Comparing Translation Quality: Google Translate Vs DeepL for Foreign Language To English. *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, 11(3), 1147–1171. <https://doi.org/10.47668/edusaintek.v11i3.1264>
- Chandra, S. O., & Yuyun, I. (2018). the Use of Google Translate in Efl Essay Writing. *LLT Journal: A Journal on Language and Language Teaching*, 21(2), 228–238. <https://doi.org/doi.org/10.24071/lt.2018.210212>
- Chen, A., Song, Y., Zhu, W., Chen, K., Yang, M., Zhao, T., & Zhang, M. (2025). *Evaluating o1-Like LLMs: Unlocking Reasoning for Translation through Comprehensive Analysis*. <http://arxiv.org/abs/2502.11544>
- English First. (2024). Ef Epi. In *EF English Proficiency Index*. <https://www.ef.com/wwen/epi/>
- Fadila, L., Dalimunthe, S. A., & Siagian, N. (2024). Strategies to Improve English Reading Comprehension. *Fonologi: Jurnal Ilmuan Bahasa Dan Sastra Inggris*, 2(3), 16–25. <https://doi.org/https://doi.org/10.61132/fonologi.v2i2.774>
- Giawa, D. (2022). Teacher Strategies in Teaching Reading Comprehension At the Eighth Grade Student of Junior High School 1 Umbunasi. *FAGURU:Jurnal Ilmiah Mahasiswa Keguruan*, 1(2), 8–16. <https://doi.org/https://doi.org/10.57094/faguru.v1i2.504>
- Kim, H., & Cha, Y. (2023). *The Role of AI Translators on Reading Comprehension*. 23, 38–58. <https://doi.org/10.15738/kjell.23..202301.38>
- Mauliska, N., & D'Angelo, J. F. (2024). The Importance of Learning English at School. *Interling : International Journal of English Language Teaching, Literature and Linguistics*, 2(2), 53–57. <https://doi.org/https://doi.org/10.55210/interling.v2i2.1793>
- Nurdiana, N., & Amelia, R. (2017). *Interpretive Reading*. Kreasi Edukasi. https://repository.uin-suska.ac.id/49373/1/buku_lengkap_Interpretive_reading.pdf
- OECD. (2023). PISA 2022 Results Factsheets Indonesia. In *OECD (Organisation for Economic Co-operation and Development) Publication*. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/indonesia_c2e1ae0e-en.html
- Prihatini, S. O. (2020). an Analysis of Students' Difficulties in Reading Comprehension at Sma Negeri 1 Sukodadi Lamongan. *E-Link Journal*, 7(1), 21–29.

<https://doi.org/https://doi.org/10.30736/ej.v7i1.261>

Rahmawati, Y., Khatimah, K., Rachman, D., & Sunarti, S. (2024). Students' Reading Comprehension of English Textual Literacy Among Zillennial Generation. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, 11(1), 82–95. <https://doi.org/https://doi.org/10.33394/jo-elt.v11i1.10985>

Sahmal, J. A., Sutaryo, Jusnita, N., & Dahlan, S. (2022). Students' Reading Comprehension in Reading English Textbook. *Jurnal Bilingual*, 12(1), 57–62. <https://doi.org/https://doi.org/10.33387/j.bilingual.v12i1.4874>

Weigang, L., & Brom, P. C. (2025). *The Paradox of Poetic Intent in Back-Translation: Evaluating the Quality of Large Language Models in Chinese Translation*. 1–25. <http://arxiv.org/abs/2504.16286>