



E-ISSN: 2789-1615  
P-ISSN: 2789-1607  
[www.educationjournal.info](http://www.educationjournal.info)  
Impact Factor: RJIF 5.93  
IJLE 2025; 5(2): 413-417  
Received: 23-07-2025  
Accepted: 28-08-2025

**Dan Kramer**  
Dean, Institute of Biblical  
Translation, Barclay College,  
Haviland, Kansas, USA

## Using artificial intelligence to check biblical texts: A report of a preliminary trial

**Dan Kramer**

**DOI:** <https://www.doi.org/10.22271/27891607.2025.v5.i2e.359>

### Abstract

This article details the findings of a preliminary trial of using Artificial Intelligence (AI) to check a simplified version of an English biblical text against a published version of an English biblical text that is on the same reading level. In the trial, the *Children's Bible*, a simplified English text that was developed by the Our Daily Bread Ministries Literacy Program, was checked against an easy-to-read *New International Reader's Version of the Bible (NIRV)*. The two texts were compared according to Accuracy, Clarity, and Naturalness, and were also evaluated on the degree to which the texts displayed similar transfer of meaning with regard to theological aspects. Findings revealed that an AI system (chatGPT) could provide a fast and efficient manner of evaluating the texts according to all aspects (Accuracy, Clarity, and Naturalness). When using AI as a preliminary text checking tool, human consultants could pinpoint areas where any corrections or other textual modifications might be necessary and, in discussion with a partner or larger group, determine the necessary corrections. When published, the *Children's Bible* will be used by the Our Daily Bread Ministries Literacy Program. It will also serve as a freely available source text for Bible translation into indigenous languages by translators who have a working knowledge of English, but who do not have access to the Bible's original languages of Hebrew, Aramaic and Greek.

**Keywords:** Artificial intelligence, comparison of source and target texts, accuracy, clarity, naturalness, consultant checking, transfer of meaning, chatGPT

### Introduction

Artificial Intelligence, or AI, and its myriad applications to industry, medicine, education, and just about every other facet of society have only begun to be realized. While some fear the technology as something that will infringe upon human creativity or replace humans in the workforce, many see its positive uses as well, especially when it is applied to boring and/or repetitive tasks that most people dread doing.

One very time-consuming and labor intensive task is that of the quality checking of translated Scripture, which is often referred to as consultant checking. Using traditional methods, consultant checking of even one book of the Bible (e.g., one of the four gospels, or one of Paul's epistles) can consume tremendous amounts of human and fiscal resources, and can take up to a few weeks before a biblical text is ready to be published.

Through the years, most Bible translation consultants have come to agree that when checking a new translation with a source text, at least three important aspects should be effectively addressed: accuracy, clarity, and naturalness. Accuracy pertains to whether the target text includes all of the elements that are found in the source text and that nothing is added or distorted in any way. Clarity relates to whether the meaning that is being transferred is clear and understandable without the recipient of the message having to exert an undue amount of cognitive load. Naturalness, which is a cultural aspect of translation, specifies whether a text sounds like it was produced by a native speaker and is appropriate for the setting for which it is intended. When translated biblical texts are checked verse-by-verse with a source text, and accuracy, clarity, and naturalness are assessed at every point along the way, it is quite understandable that the entire process can take extreme amounts of time. Then, when one factors in elements like limited human attention spans and human fatigue, the likelihood of errors creeping into the target text increases dramatically.

Concerns such as the above prompted our researchers to examine whether AI might serve an effective role in comparing two texts by providing a helpful means of evaluating a target text for accuracy, clarity, and naturalness, when compared with a source text. Because all of the researchers who participated in this study were experienced Bible translation consultants, an effective assessment of the AI evaluation could be readily checked by the "humans in the

**Correspondence Author;**  
**Dan Kramer**  
Dean, Institute of Biblical  
Translation, Barclay College,  
Haviland, Kansas, USA

loop.” This would enable our researchers to determine whether AI might be a helpful tool in facilitating the process of Bible translation consultant checking. If so, this “symbiotic” union between AI and humans could greatly accelerate a process that has been painfully slow, while contributing to the production of biblical texts that are of high quality from the standpoints of accuracy, clarity, and naturalness.

### Literature Review

The present study represents some of the first efforts to use an AI model (chatGPT) to analyze two biblical texts and evaluate them according to accuracy, clarity, and naturalness. While AI models have not yet been used to consultant check translated biblical texts in preparation for publication, AI models have been used to analyze biblical texts over the past 10 years or so to address concerns such as authorship and other academic topics such as textual criticism.

Lima, Omar, Avansi, and de Castro (2025) <sup>[8]</sup> conducted a systematic review of AI and machine-learning applications in biblical text research, which summarized 85 papers across seven categories. Areas of exploration that were relevant to the present study included: analysis of texts to determine authorship, using AI handwriting analysis to identify distinct scribal hands in a manuscript, the use of AI models to explore literary themes, and discovering limitations of neural translation models in the effective handling of semantic nuance in biblical narratives.

Levy, Ganor, and Beeri (2016) <sup>[7]</sup> used machine-based handwriting analysis to shed light on the composition of biblical texts and Hebrew inscriptions. Rosa (2025) <sup>[11]</sup> used “stylometric clustering” to explore authorship patterns within the corpora of the Pauline letters in the New Testament. Faigenbaum, Golovin, Kipnis, *et al.* (2024) used statistical word-frequency analysis to explore authorship attribution within the Hebrew Bible.

Popović, Dhali, Schomaker, *et al.* (2025) <sup>[10]</sup> introduced an AI model “Enoch,” which provided paleographic dating for some of the Dead Sea Scrolls, while conducting handwriting analysis to identify unique scribal hands in the Great Isaiah Scroll.

AI has also provided a great tool for analyzing texts for a variety of academic pursuits within the field of Biblical Studies. Zhao and Liu (2018) <sup>[13]</sup> developed a neural model to explore machine comprehension of scripture. Zhao and Liu’s work is particularly relevant to the present study since it helps to provide a foundation for the ability of an AI model to “understand” a text in preparation for determining whether meaning has been accurately transferred from a source text to a target text.

Other studies which have used AI models to analyze texts for semantic and cross-lingual alignment include Kauhanen and Kalavainen (2020) <sup>[5]</sup>, which described computational tagging of variant apparatus data in the Göttingen Septuagint apparatus; Jones, *et al.* (2022) <sup>[4]</sup>, which examined machine learning algorithms for manuscript variant weighting and reconstruction of biblical texts; and, Dörpinghaus (2024), which automated cross-lingual alignment in parallel multilingual Bible corpora.

AI has also been used to identify literary themes in biblical literature. Studies such as Kong, Kang, Park, and Park (2025) <sup>[6]</sup> demonstrated how embedding-based AI models can be used to detect thematic threads, patterns, and style

shifts across biblical translations. A related study, by McGovern, Sirin, and Lippincott (2025) <sup>[9]</sup>, used a model trained on neural embeddings to identify the literary device known as chiasm in biblical texts. This is particularly helpful in theme identification, as well as in determining the likelihood that analyzed texts might have been produced by a unique author. These two studies are relevant in that they both have demonstrated the ability of AI to analyze texts in a way that can inform the present study’s identification of theological themes and the transference of their meaning across texts.

While these early AI analyses of biblical texts have demonstrated a high level of usefulness, it must also be acknowledged that AI may not always be reliable to produce outcomes that can be readily accepted without further verification by specialist human judgment, or what has sometimes been referred to as “the human in the loop.” For example, Rosa (2025) <sup>[11]</sup> cautioned against using a “computational stylometric” model exclusively in concluding whether authorship can be determined with certainty for a text under analysis. Similarly, Bénel, Falip, and Lacour (2024) <sup>[1]</sup> demonstrated the limitations of the use of AI in biblical contexts in their study entitled, “When Abel Kills Cain: What Machine Translation Cannot Capture.”

With information on the strengths and weaknesses of AI as an analytical tool for biblical texts, the authors of this study set out to determine if the freely available chatGPT model might be a valuable tool in analyzing a simple English biblical text alongside a published biblical text for the universally accepted Bible translation criteria of accuracy, clarity, and naturalness.

### Research Problem and Questions

Our Daily Bread Ministries (ODBM) has developed an easy-reading *Children’s Bible* to be used in its global Literacy Program. This Bible will also be used as a source text for future Bible translations into indigenous languages in Africa and throughout the world. Having an easy-to-read source text will enable Bible translators who do not have access to the Bible’s original languages (Hebrew, Aramaic, and Greek), but who do have a basic command of English, to accomplish their translations. In order for ODBM to effectively use its Bible in literacy training and translation, care must be taken that the text of the *Children’s Bible* is accurate, clear, and natural, when compared to a standard published text that has been based on the original Biblical languages.

To assure that a biblical text meets the three-fold quality criteria (accuracy, clarity, and naturalness), the text must be carefully checked by qualified persons who have been trained for this task and who have had years of experience checking biblical texts in preparation for publishing. In the Bible Translation world, these are known as Bible Translation consultants. Bible translation agencies often report that the checking process can take up to several years or more, if a consultant is even available to do the work. Because of this checking “bottle neck” in Bible translation, a method is needed that will speed up the process so that these translations can be received by the people for whom they are intended. Yet, with the increased speed of doing the checking process, the quality of the text must also be assured.

Based on the use of Artificial Intelligence as it is currently being applied in numerous areas of Biblical Studies, our

researchers attempted to determine whether AI could be used to evaluate a target text against a standard published biblical text. Because ODBM's *Children's Bible* was developed to be used as a resource in teaching children to read, it was determined that it should be evaluated alongside a standard Bible that was written on a similar reading level. Once a suitable standard text could be identified, the team of researchers could use AI (chatGPT) to evaluate the target text according to the desired quality criteria of accuracy, clarity, and naturalness. Further evaluation could be conducted to assure that theological aspects and themes in the source text were transferred appropriately in the target text.

### Methodology

A trial was organized and led by Our Daily Bread Ministries' department of strategic initiatives in cooperation with the Biblical Translation Institute of Barclay College, Haviland, Kansas. All members of the research team were trained as Bible Translation consultants with several years experience in guiding translation projects and checking the biblical texts for publication. Three of the team members were university professors.

### Protocol for Checking Children's Bible using AI (ChatGPT)

**Place of Text Checking Trial:** Corinth Baptist Church - Elizabeth City, NC (10/17 - 10/18/2025)

**Cooperating Partners:** Our Daily Bread Ministries (Grand Rapids, MI) and the Biblical Translation Institute of Barclay College (Haviland, KS)

1. Ask chatGPT to determine which published version of the Bible most closely matches our text (*Children's Bible*). This will be our source text for comparison. In our case, the closest match was determined to be the *New International Reader's Version (NirV)*.
2. Load CB text portion (chapter or portion of chapter) into chatGPT. Ask chatGPT to verify our CB text against the selected Source text according to the following criteria:
  - a. **Accuracy:** Is there anything added? Is anything left out? Are there any distortions?
  - b. **Clarity:** Does the text use simple words and short-styled sentences? Is the text easy to understand?
  - c. **Naturalness:** Does the text use conversational language? Is the register of the text informal, yet culturally appropriate?
  - d. **Consistency:** Capitalization, word choice (to a pre-determined degree). Questions to explore with regard to consistency: Should the consistency criterion be assessed with A,C, and N, or as a separate step? Must all key words be translated identically, or are near synonyms acceptable? In our checking, we only used Accuracy, Clarity, and Naturalness; however, chatGPT did offer some comments on consistency.
3. Ask chatGPT to color code the degree to which the

target text (CB) matches the source text (NirV), using green for excellent, yellow for good, but with some discrepancies, and red for significant problems.

4. Ask chatGPT to offer suggestions for more closely aligning the CB with the style of source text. ChatGPT might also offer suggestions without asking.
5. Make changes according to chatGPT suggestions and track changes in Pages or Word. Be sure author identification is turned on and unique author color is selected.

### Results

In order to determine which published biblical text was similar to the target (*Children's Bible*) and on the same reading level, a prompt was submitted to chatGPT. The response, which was aligned with our previous research was that the *New International Reader's Version (NirV)* of the Bible was similar to the *Children's Bible* (CB) and was approximately on the same reading level (4th to 5th grade). Next a portion of our CB text (Gospel of John, chapter one) was loaded into chatGPT and a prompt was submitted, asking chatGPT to evaluate our target text against the NirV text for the same section of biblical text. ChatGPT was asked to evaluate the target text according to accuracy, clarity, and naturalness. As our prompts were fine-tuned to produce more readable results, our researchers asked chatGPT to color code the degree to which the target text (CB) matches the source text (NirV), using green for excellent, yellow for good, but with some discrepancies, and red for significant problems that needed to be addressed by the group. It was determined that, with regard to consistency between the source and target texts, word-for-word translations (formal equivalence) would not be necessary to assess accurate transfer of meaning of theological aspects (key words), but that near-synonyms would be accepted in the goal of achieving dynamic equivalence between the text. After numerous attempts to develop effective prompts that delivered desired responses, the researchers decided upon slight variations of the following prompt as the most usable in obtaining the best responses from chatGPT:

**Prompt:** Compare our *Children's Bible* (CB) text of John chapter one to the New International Reader's Version of that same text. Evaluate the CB text according to accuracy, clarity, and naturalness for each verse. Show findings on a screen-friendly chart, using color-coded icons (green for excellent, yellow for good, but some possible concerns, and red for not acceptable).

**Table1:** Summary of chatGPT Evaluation of Children's Bible (John chapter one)

| Category        | Average Rating | Overall Assessment                                |
|-----------------|----------------|---------------------------------------------------|
| **Accuracy**    | □Excellent     | Highly faithful to source.                        |
| **Clarity**     | □Excellent     | Very readable for children and general audiences. |
| **Naturalness** | □Excellent     | Smooth, idiomatic English throughout.             |

**Table 2:** Summary of chatGPT Evaluation of Children's Bible (John chapter one, verse by verse analysis with notes for possible discussion)

| Verse | Accuracy | Clarity | Naturalness | Notes                                                              |
|-------|----------|---------|-------------|--------------------------------------------------------------------|
| 1     | ☐        | ☐       | ☐           | Faithful rendering; reads smoothly and naturally.                  |
| 2     | ☐        | ☐       | ☐           | "They" slightly obscures Trinitarian nuance ("He").                |
| 3     | ☐        | ☐       | ☐           | Very close to NIV; natural and age-appropriate.                    |
| 4     | ☐        | ☐       | ☐           | Excellent paraphrase preserving key theological meaning.           |
| 5     | ☐        | ☐       | ☐           | "Did not put it out" captures idea clearly and naturally.          |
| 6     | ☐        | ☐       | ☐           | Faithful and simple.                                               |
| 7     | ☐        | ☐       | ☐           | Clear and accurate; "so that all may believe" natural.             |
| 8     | ☐        | ☐       | ☐           | Matches NIV conceptually and linguistically.                       |
| 9     | ☐        | ☐       | ☐           | Excellent; preserves parallel to "true Light."                     |
| 10    | ☐        | ☐       | ☐           | "People did not know Him" fits NIV's "did not recognize."          |
| 11    | ☐        | ☐       | ☐           | Clear and faithful.                                                |
| 12    | ☐        | ☐       | ☐           | "Choice to become" adds interpretive nuance not in NIV.            |
| 13    | ☐        | ☐       | ☐           | Faithful summary of NIV phrasing.                                  |
| 14    | ☐        | ☐       | ☐           | Smooth, age-appropriate, and accurate.                             |
| 15    | ☐        | ☐       | ☐           | Well adapted quotation; clear attribution.                         |
| 16    | ☐        | ☐       | ☐           | Captures "grace upon grace."                                       |
| 17    | ☐        | ☐       | ☐           | Matches NIV precisely in meaning.                                  |
| 18    | ☐        | ☐       | ☐           | Excellent summary of theological concept.                          |
| 19    | ☐        | ☐       | ☐           | Correct narrative setup.                                           |
| 20    | ☐        | ☐       | ☐           | Matches NIV; simple and direct.                                    |
| 21    | ☐        | ☐       | ☐           | Faithful sequence; good child readability.                         |
| 22    | ☐        | ☐       | ☐           | Excellent paraphrase maintaining clarity.                          |
| 23    | ☐        | ☐       | ☐           | Accurately captures Isaiah reference.                              |
| 24-25 | ☐        | ☐       | ☐           | Smooth and accurate; combined naturally.                           |
| 26-27 | ☐        | ☐       | ☐           | Faithful; wording simple and strong.                               |
| 28    | ☐        | ☐       | ☐           | Perfect alignment with NIV geography and tone.                     |
| 29    | ☐        | ☐       | ☐           | "Washes sin away" paraphrases "takes away," acceptable but softer. |
| 30    | ☐        | ☐       | ☐           | Clear and accurate.                                                |
| 31-33 | ☐        | ☐       | ☐           | Matches sequence and meaning closely.                              |
| 34    | ☐        | ☐       | ☐           | Faithful and well phrased.                                         |
| 35-37 | ☐        | ☐       | ☐           | Smooth retelling; retains dialogue accuracy.                       |
| 38-39 | ☐        | ☐       | ☐           | Clear and natural; no issues.                                      |
| 40-42 | ☐        | ☐       | ☐           | Excellent narrative cohesion.                                      |
| 43-44 | ☐        | ☐       | ☐           | Simple and faithful.                                               |
| 45-46 | ☐        | ☐       | ☐           | Faithful rendering; idiomatic and natural.                         |
| 47-48 | ☐        | ☐       | ☐           | Captures exchange naturally.                                       |
| 49    | ☐        | ☐       | ☐           | Accurate theology and tone.                                        |
| 50-51 | ☐        | ☐       | ☐           | Smooth ending; faithful to NIV phrasing.                           |

As can be seen from the above charts, chatGPT's findings indicate that John chapter one of the Children's Bible is faithful to the source text (NIV) when evaluated according to accuracy, clarity, and naturalness. In fact, chatGPT alerted our researchers to only three areas of concern (verse 2 - clarity; verse 12 - naturalness; verse 29 - naturalness). All other verses scored "excellent" for the three criteria under investigation.

Because the chatGPT evaluation of John chapter one of the Children's Bible was the first attempt during the trial, the researchers decided to verify the evaluation of all 52 verses in the chapter, not just the 3 where possible concerns were highlighted. In some cases, chatGPT's "excellent" ratings were challenged and changes were made in the *Children's Bible* text. In other cases, where chatGPT highlighted potential concerns, it was determined that the target text was acceptable without any changes.

A similar procedure was followed by our researchers throughout the course of the two-day trial. Because of the speed of using chatGPT as a preliminary text checking tool, our researchers were able to check the entire Gospel of John in just a few hours. During the course of the trial, they also checked several other New Testament books in the *Children's Bible*.

## Discussion

During the course of the weekend trial for checking texts from the *Children's Bible*, our researchers used the freely available version of chatGPT, which was accessed at chatGPT.com. The speed with which chatGPT was able to deliver resulted was quite impressive. In fact, early on in the trial, our researchers seemed reluctant to accept even the "excellent" verse ratings without investigating further. Because some changes were made to verses that were rated "excellent," our researchers continued to verify all of the results, and not just the ones that were highlighted as concerning. As our researchers became more confident in using chatGPT's ratings as valid, fewer such investigations of "excellent" ratings were conducted, and the focus shifted more to verses that were marked as potentially problematic. This enabled the researchers to work at a much faster pace, which resulted in a large amount of text being checked and approved for publication.

One concern that was noted during the trial was the inability of chatGPT to provide consistent responses even when given identical prompts. The researchers also discovered that they were sometimes getting different responses from chatGPT when everyone was using the same prompt. While this did provide some level of frustration, especially early on in the trial, the problem was easily addressed as

researchers quickly learned to vary the content of their prompts until they received the responses that they desired. For example, sometimes the charts produced by chatGPT could not be read easily on the screens that were being used by the researchers. This simply required the submission of an additional prompt or two until the desired table format was achieved. At other times, chatGPT would only provide a paragraph summary of its evaluation, which was not as usable as a color-coded chart. Again, this required additional prompts to secure the desired style of responses. These concerns characterized some of our researcher's early attempts during the trial, but were easily dealt with as they gained more experience in using chatGPT.

### Future Research Directions

Several of the researchers who participated in this trial have experience in both traditional consultant checking as well as more efficient methods of Bible translation and checking such as MAST (Mobilized Assistance Supporting Translation), which was developed by some of the present researchers (Foster, 2023) <sup>[3]</sup>. All who participated in this trial were impressed by the potential of using AI to greatly speed up the process of checking and to assure that the benchmarks of accuracy, clarity, and naturalness were met. As a result of this trial, it seems that AI, in the form of chatGPT, can be a valuable tool in effectively and quickly performing at least a preliminary quality check of biblical texts, especially when qualified experts are also participating in the work.

Future research in using other AI tools to analyze biblical and other types of texts is needed. Of particular interest might be whether paid versions of chatGPT and other platforms such as Microsoft Copilot, Google Gemini, or Apple Intelligence would perform as well or better than chatGPT to analyze and evaluate texts. Since the purpose of this trial was to evaluate the potential of AI to assist in bringing Our Daily Bread Ministry's *Children's Bible* to the field as soon as possible, chatGPT performed at a level that met or exceeded the expectation of our researchers.

### Acknowledgments

Heartfelt thanks and appreciation are extended to the research team who, in addition to the authors, participated in this project. These were Steve Cunningham, Mike Hatfield, and Cheryl Luton. A special word of thanks to Mary Cunningham and Cheryl Luton for meals and snacks that were provided to the team. An additional word of appreciation is extended to Pastor Greg L. Carr and Corinth Baptist Church, Elizabeth City, North Carolina, for the use of its facilities during the project, and for providing lunch on the second day of the trial.

### References

1. Bénel A, Falip J, Lacour P. When Abel kills Cain: What machine translation cannot capture. arXiv:2404.04279. 2024.
2. Faigenbaum-Golovin S, Kipnis A, Bühler A, Piasetzky E, Römer T. Critical biblical studies via word frequency analysis: unveiling text authorship. arXiv:2410.19883. 2024.
3. Foster M. Transformation in Bible Translation: Accelerated Methods Yield Similar Results as Benchmark Processes. 2023.
4. Jones M, Romano F, Mohd A. Using artificial intelligence to reconstruct the text of the New Testament. In: Proceedings of the 2022 7th International Conference on Machine Learning Technologies. 2022.
5. Kauhanen T, Kalavainen HPK. Automatic Tagging of the Göttingen Septuagint Apparatus. TC J Bibl Text Crit. 2020;25:145-147.
6. Kong I, Kang S, Park Y, Park S. Style Extraction on Text Embeddings Using VAE and Parallel Dataset. arXiv:2502.08668. 2025.
7. Levy T, Ganor H, Beeri R. Algorithmic handwriting analysis of Judah's military correspondence sheds light on composition of biblical texts. J Archaeol Sci. 2016;67:32-37.
8. Lima BC, Omar N, Avansi I, de Castro LN. Artificial Intelligence Applied to the Analysis of Biblical Scriptures: A Systematic Review. Analytics. 2025;4(2):13.
9. McGovern H, Sirin H, Lippincott T. Computational Discovery of Chiasmus in Ancient Religious Text. In: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies. Vol 2. 2025;p. 154-160.
10. Popović M, *et al.* Dating ancient manuscripts using radiocarbon and AI-based writing style analysis. arXiv:2407.12013v2. 2025.
11. Rosa A. Computational Stylometrics and the Pauline Corpus: Limits in Authorship Attribution. Religions. 2025;16(10):1264.
12. The Holy Bible, New International Reader's Version. Zondervan; 1996.
13. Zhao HJ, Liu J. Finding Answers from the Word of God: Domain Adaptation for Neural Networks in Biblical Question Answering. arXiv:1810.12118. 2018.