International Journal of Transformative Multidisciplinary Studies
ISSN Print: 3116-3009 | ISSN Online: 3116-3017
Driving Transformative Impact Across Disciplines

Peer Review Report Form

Manuscript Title:	Enhancing the Writing Proficiency of Non-English Major Students Using Artificial Intelligence (AI) Language Tools
Type of Article:	Original Article

FIRST ROUND

Editor

General Comments and Recommendations
The reviewers acknowledge the significance of your manuscript and recommend several revisions to further improve its clarity, coherence, organization, and overall scholarly quality. In addition, you are advised to thoroughly review and validate all references included in the manuscript. Please ensure that all cited sources are accurate, accessible, and retrievable through Google Scholar to uphold the credibility and reliability of the study. Additionally, make sure to include the DOI or website link for every reference whenever available to facilitate verification and accessibility of sources. Furthermore, only sources that are cited within the manuscript should appear in the References section, and all entries listed in the References section must also be cited in the text. For example, if 20 unique sources are cited in the manuscript, then only those 20 sources should be included in the References section. Please also note that repeated citations of the same source within the text are counted as a single reference entry; therefore, the total number of references may be lower than the total number of in-text citations.

Reviewer 1

General Comments and Recommendations
The paper is timely and useful because AI writing tools are now really part of students' academic work. The study also has a clear intention to help non-English major students improve their writing, not just to measure scores. However, the main issue that needs serious improvement is the research design and interpretation of causality. As written, the paper strongly claims that AI language tools "enhanced" or "caused" improvement in writing proficiency, but the design used is only a one-group pretest-posttest design with no control/comparison group. Because of this, it is difficult to separate the effect of AI from other possible explanations such as practice effect, repeated exposure to the same task, familiarity with the rubric, teacher/researcher guidance, or the natural improvement from revising an earlier draft. This is important because the whole paper depends on the claim that AI was the reason for the improvement. The manuscript itself states that only 77 students from the original 169 proceeded after selecting those at Beginning to Emerging levels, and these same students revised their own essays using AI feedback, then their pretest and posttest scores were compared.

International Journal of Transformative Multidisciplinary Studies
ISSN Print: 3116-3009 | ISSN Online: 3116-3017
Driving Transformative Impact Across Disciplines

Specific Comments and Recommendations
I suggest that the authors be more careful in presenting the findings. The study can still be valuable, but the conclusion should be softened. Instead of saying that AI language tools "significantly enhance" writing proficiency, it may be more accurate to say that students showed significant improvement after AI-assisted revision, but this improvement cannot be attributed to AI alone because of the absence of a control group. The authors also need to explain more clearly how the scoring was protected from bias, especially since the post-test was a revised version of the pre-test essay and the adviser moderated the scoring. It would help if the authors add inter-rater reliability, blind scoring procedure, and clearer control of the AI intervention. Also, the interpretation of statistical results should be checked: the methodology says paired-samples t-test, but Table 3 reports an F value and partial eta squared, which looks more like ANOVA output. This inconsistency weakens confidence in the analysis and should be corrected.

Please indicate your recommendation by checking the appropriate box below.

Decision	
	Accept the manuscript for publication.
/	Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.
	Reject the manuscript, as it is not suitable for publication.

Reviewer 2

General Comments and Recommendations
Thank you for the opportunity to review this manuscript. The researcher is commended for tackling a very timely and relevant issue, which is the integration of AI tools in education, specifically for non-English major students. The creation of the Project A.I.-V.O.I.C.E. intervention program shows a good effort to produce a practical, localized output from the research.
Specific Comments and Recommendations
The study utilizes a single-group pretest-posttest design. This is a very weak experimental design to establish cause and effect. The students wrote a first draft, and then they were given 1 hour to use AI and revise it. Naturally, a revised second draft will almost always score higher than a rushed first draft, even without AI. Because there is no control group (a group that spent 1 hour revising without AI assistance), you cannot confidently claim that the AI is the main reason for the significant improvement. The improvement could just be the result of having an extra hour to edit.
The methodology states that after the pretest, only the students who scored at the "Beginning" to "Emerging" levels were selected for the post-test. By purposely removing the high-performing students and only testing the low scorers, the study is vulnerable to a statistical phenomenon called "regression to the mean." This means low scorers on a first test are mathematically highly likely to score higher on a second test just by chance. This severely biases your paired-samples t-test results.

International Journal of Transformative Multidisciplinary Studies
ISSN Print: 3116-3009 | ISSN Online: 3116-3017
Driving Transformative Impact Across Disciplines

The title and research questions claim to measure the enhancement of "writing proficiency". However, the intervention was only a 1-hour session using ChatGPT or Gemini. One hour of using a tool to edit a single essay does not measure a student's actual long-term cognitive writing proficiency; it only measures the quality of the AI-edited output. The researcher even admits in the limitations that the study did not measure long-term cognitive retention. The claims in the abstract and conclusion need to be toned down. You are measuring the enhancement of the written output with AI assistance, not the permanent enhancement of the student's personal proficiency.
The proposed framework is very extensive and has five detailed phases. While it looks impressive, this framework was not actually tested in the methodology. The respondents just used normal AI platforms for an hour to get rubric-aligned feedback. Proposing a massive institutional framework based on a 1-hour generic AI editing session is a bit of a stretch. You should clarify that this framework is purely theoretical at this stage and requires separate empirical testing, which you briefly mentioned in your recommendations but should be made clearer earlier on.

Please indicate your recommendation by checking the appropriate box below.

	Decision
	Accept the manuscript for publication.
/	Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.
	Reject the manuscript, as it is not suitable for publication.

SECOND ROUND

Decision	**Editor**	**Reviewer 1**	**Reviewer 2**
Accept the manuscript for publication.	/	/	/
Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.			
Reject the manuscript, as it is not suitable for publication.			