

REVIEWER'S REPORT

Manuscript No.: IJAR-57600

Title: Assessing AI-Supported Academic Writing: Threshold Performance, Traceability, and Pedagogical Mediation in a First-Year University Course,

Recommendation:

Accept as it is

Accept after minor revision.....

Accept after major revision

Do not accept (*Reasons below*)

Rating	Excel.	Good	Fair	Poor
Originality		√		
Techn. Quality			√	
Clarity			√	
Significance		√		

Reviewer's ID: JPR- 212

Detailed Reviewer's Report

Summary of the Manuscript

This mixed-methods study compares three pedagogical conditions of academic writing in a first-year university Spanish course: AI-integrated in-situ writing (n=11), non-AI in-situ writing (n=13), and non-AI take-home writing (n=11). Using institutional rubrics aligned with Effective Communication, Information Literacy, and Critical Thinking, plus qualitative analysis of 12 AI writing logs, the authors find that while the take-home condition produced the highest overall scores, the AI-integrated condition generated more traceable evidence of student decision-making. The study argues that the educational value of AI in writing depends on pedagogical mediation and traceability, not merely on the final product.

Major Concerns (Reasons for "Accept after major revision")

1. Fundamental Confounding of Variables Undermines Comparative Claims The three conditions differ simultaneously across multiple dimensions: AI use (yes/no), setting (in-situ vs. take-home), time available, documentation requirements, and access to revision support. The authors acknowledge this limitation but then proceed to make comparative claims (e.g., "AI-integrated condition generated the most pedagogically traceable evidence") that are not logically supported. The traceability advantage of the AI condition is entirely due to the *requirement to complete writing logs*, not to AI itself. A non-AI in-situ condition with equivalent log documentation would be the necessary control, but it does not exist. This is a fatal design flaw.

2. The Central Claim Is a Tautology Disguised as a Finding The paper concludes that AI-supported writing is valuable "when it makes student decision-making visible." However, the study operationalized "value" as *having a writing log*. The AI-integrated condition produced more traceable evidence only because logs were mandated. The take-home condition produced no logs. Therefore, the finding reduces to: requiring documentation produces documentation. No meaningful comparison of AI's pedagogical contribution is possible.

3. Critical Thinking Findings Are Overinterpreted The AI-integrated condition failed to meet the threshold for Critical Thinking (65% vs. 77% for take-home). The authors suggest AI "may support expression before deepening ideas," but the data do not rule out simpler explanations: (a) the AI condition had less time (in-class), (b) the AI condition had a smaller sample (n=11), (c) the AI tool was

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

used in a highly constrained, teacher-prompted manner. Without controlling for time and task difficulty, claims about AI's differential effect on Critical Thinking are speculative.

4. Sample Size and Statistical Treatment With $n=11$, 13, and 11 across conditions, descriptive percentages (e.g., 81.82% reaching threshold) have wide confidence intervals. No inferential statistics are reported (e.g., no chi-square, t-test, or ANOVA), yet the discussion repeatedly uses comparative language ("higher," "more balanced," "strongest"). The lack of statistical testing, combined with the small sample, means the reported differences may be due to sampling variation. The authors state they "did not aim to establish causal effects," but the manuscript's structure and conclusions repeatedly imply comparative validity.

5. Missing Critical Information for Replicability

- The exact AI prompt provided to students is not included.
- The institutional rubric criteria (8-point scale) are described but not reproduced in full.
- The inter-rater reliability of rubric scoring is not reported (only a passing reference to Koo & Li, 2016 in references, but no coefficients are given).
- The qualitative analysis does not report coding reliability, number of coders, or resolution of disagreements.
- The AI writing log template is not provided as an appendix.

6. Overclaiming Institutional and Practical Implications The authors argue that "institutions should consider documenting task conditions as part of the assessment record." This is reasonable, but the study provides no evidence about the feasibility, faculty acceptance, student compliance, or scalability of AI writing logs in larger courses or across disciplines. The recommendation is asserted without any implementation data.

Minor Concerns

- **Redundancy:** The conceptual framework (threshold performance, traceability, pedagogical mediation) is restated nearly verbatim in the discussion section, lengthening the paper without adding insight.
- **Unsupported Terminology:** The phrase "metaevidence" is introduced but not defined or operationalized.
- **Selective Citation:** Claims about equity assessment (Montenegro & Jankowski, 2020) and authentic assessment (Jankowski et al., 2018) are cited appropriately, but the study does not actually measure equity outcomes (e.g., differential performance by prior academic preparation, language background, or access to technology at home).
- **Missing Negative Cases:** The qualitative analysis reports that students sometimes rejected AI suggestions, but no examples are provided of students who uncritically accepted poor suggestions. The absence of disconfirming evidence suggests positive selection bias.
- **Spanish-Language Context Underutilized:** The sociolinguistic context of Puerto Rico is introduced in the introduction but never analyzed in the results. No comparison is made between Spanish-specific writing features and English-medium AI tool behavior.

Ethical and Transparency Concerns

The study states that "writing tasks...were used for formative and institutional interpretation, not to assign summative consequences." However, students in different conditions received systematically different instructional support. It is not clear whether students were informed of the different conditions prior to consent, or whether the take-home group knew that the AI group received structured revision support. This raises questions about equitable treatment within a single course.

Suggested Revisions for Resubmission (If Major Changes Were Possible)

The paper would require fundamental redesign to be publishable. Minimal requirements for a revised submission would include:

1. A fourth condition: non-AI in-situ *with writing logs* to isolate the effect of documentation from the effect of AI.

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

2. Statistical testing with appropriate corrections for small samples (or a clear statement that only exploratory, hypothesis-generating claims are made, with no comparative conclusions).
3. Full rubric, AI prompt, and log template in appendices.
4. Inter-rater reliability coefficients for both rubric scores and qualitative coding.
5. A revised abstract and conclusion that do not overstate comparative findings.

Given these requirements, the current manuscript cannot be accepted even with major revision, as the design confound is fundamental.

Overall Assessment: The paper addresses an important and timely topic. The conceptual framework (threshold performance, traceability, pedagogical mediation) is thoughtful. The qualitative excerpts from AI logs are genuinely interesting and suggest that writing logs may be a useful pedagogical tool. However, the comparative design does not support the central claims about AI's unique contribution. Without a non-AI logged condition, the study cannot distinguish between the effect of AI and the effect of documentation. Therefore, the manuscript is not suitable for acceptance in its current form