

Unknown Paper

AI Reviewer 1 - Comprehensive Review

This reviewer provides a thorough and balanced assessment

****Review of "WESEEIT,BUT NOONEBELIEVES US": TRUSTWORTHINESS IN COMMUNITY TESTIMONY AND DATA REPORTING DURING GAZA'S 2024-2025 RECONSTRUCTION****

\\ **** Step 1. Summary of the Paper**** This manuscript examines the credibility of community-reported infrastructure damage data from Gaza (October 2023–June 2024) using the "Genocide of the Palestinian People" dataset. The authors employ a mixed-methods approach: quantitative analysis of 255 daily records (assessing internal consistency via correlations and temporal trends) and qualitative analysis of 18 simulated interviews with engineers, volunteers, and residents. The paper claims to: (1) validate the statistical coherence of community-generated data, (2) highlight epistemic injustices in institutional reception (e.g., data dismissal or revision without consultation), and (3) propose an integrated framework combining statistical rigor and ethical principles (epistemic justice) to establish trustworthiness in humanitarian data.

\\ **** Step 2. Evaluation Criteria****

\\\ ****1. Originality / Novelty**** ****Score: 6/10**** - ****Strengths****: The focus on epistemic justice (Fricker, 2007) in humanitarian data contexts is underexplored, particularly in conflict zones like Gaza. The integration of mixed methods to bridge statistical and ethical dimensions is a moderate contribution. - ****Weaknesses****: The core premise—community data facing institutional skepticism—is well-documented in participatory GIS and crisis mapping literature (e.g., Onen et al., 2018; Smit, 2021). The "simulated interviews" approach lacks precedent and risks undermining authenticity. - ****Verdict****: Incremental rather than field-advancing. The ethical framing is valuable but not transformative.

\\\ ****2. Scientific Rigor / Methodology**** ****Score: 3/10**** - ****Critical Flaws****: - ****Simulated Interviews****: The use of "simulated" interviews (Section 4.5) is methodologically unsound. It is unclear whether these are hypothetical, AI-generated, or role-played, raising serious questions about data validity, ethical compliance, and generalizability. - ****Data Source****: The Kaggle dataset ("Genocide of the Palestinian People") is cited but not critically evaluated for biases, verification protocols, or potential manipulation. The claim of "zero missing entries" in a conflict zone is extraordinary and requires justification. - ****Statistical Analysis****: Correlation analysis

(e.g., $r=0.967$ between educational damage variables) may reflect collinearity or redundant reporting rather than true consistency. No controls for confounding factors (e.g., media attention, access constraints) are mentioned. - **Ethical Oversight**: No evidence of IRB approval or informed consent for human participants (Section 4.8 vaguely references compliance but lacks details). - **Verdict**: Fundamental methodological flaws compromise the study's validity.

\\\\ **3. Clarity \ Presentation** **Score: 5/10** - **Strengths**: The structure follows conventional sections (Introduction, Methods, Results, Discussion), and the abstract clearly outlines goals. - **Weaknesses**: - **Jargon Overuse**: Terms like "convergent concurrent triangulation" and "epistemic trust" are repeated without sufficient clarity for interdisciplinary readers. - **Ambiguity**: "Simulated interviews" is never defined. The conflation of "moral injury" with data revision is provocative but poorly substantiated. - **Figures/Tables**: Absent. Visualizations of temporal trends or correlation matrices would strengthen quantitative claims. - **Verdict**: Readable but obfuscated by undefined terms and missing visuals.

\\\\ **4. Reproducibility \ Transparency** **Score: 2/10** - **Data/Code**: The Kaggle dataset is publicly available, but analysis scripts (Python) are not shared. - **Methodological Gaps**: No interview protocols, coding manuals, or raw qualitative data are provided. The "simulated" nature of interviews precludes replication. - **Statistical Transparency**: Correlation results are reported without p-values or confidence intervals. The claim that "correlations ≥ 0.9 correspond to a 38% - **Verdict**: Critically deficient. Replication is impossible.

\\\\ **5. Significance \ Impact** **Score: 4/10** - **Potential**: The topic is urgent—community data in Gaza could inform humanitarian response and accountability. - **Limitations**: Methodological flaws undermine policy relevance. The proposed "integrated framework" is too abstract for practical implementation. - **Audience**: Experts in humanitarian ethics may find the epistemic justice discussion engaging, but data scientists will question the analytical rigor. - **Verdict**: Local relevance but unlikely to influence broader field.

\\\\ **6. Ethics \ Integrity** **Score: 3/10** - **Ethical Concerns**: - **Simulated Interviews**: If these involve human participants, the lack of detail on consent and IRB approval is unacceptable. If AI-generated, this must be explicitly stated to avoid misrepresentation. - **Neutrality**: The dataset title ("Genocide of the Palestinian People") and framing assume a contested political stance, risking bias. The paper does not address how this affects objectivity. - **Conflicts of Interest**: No declarations provided. - **Verdict**: Ethical and integrity issues are severe and unresolved.

\\\\ **Step 3. Specific Suggestions for Improvement**

\\\\ **Major Revisions Required**: 1. **Replace Simulated Interviews**: Conduct real interviews with verified participants, detailing IRB approval and consent procedures. If simulation was used for safety reasons, justify and transparently describe the process (e.g., LLM-generated responses with validation). 2. **Validate Dataset Credibility**: Include external checks (e.g., satellite imagery correlation) to verify damage reports. Acknowledge potential biases (e.g., under-reporting in northern Gaza). 3. **Strengthen Statistical Analysis**: Add controls for temporal autocorrelation, report effect sizes and uncertainty metrics, and address collinearity (e.g., variance inflation factors). 4. **Clarify Ethical Compliance**: Provide IRB approval number and consent forms in supplements.

\\\\ \\ **Minor Revisions**: - Define key terms (e.g., "epistemic trust," "moral injury") in the introduction. - Add visualizations: time-series plots, correlation matrices, theme frequency charts. - Tone down polemical language (e.g., "genocide" in dataset title) or justify its use academically. - Fix formatting errors (e.g., inconsistent reference styling, typos like "T RUSTWOR - THINESS").

\\\\ \\ **Additional Experiments/Analyses**: - Compare community data with remote sensing estimates (e.g., Holail et al., 2024) to assess convergence. - Include gender-disaggregated analysis of interview responses to explore intersectional epistemic injustice. - Conduct sensitivity analyses for correlation thresholds (e.g., how results change if $r \geq 0.8$ vs. $r \geq 0.9$).

—

\\\\ \\ ** Step 4. Final Decision \ Justification**

\\\\ \\ **Overall Score: 3/10** The paper addresses a timely and morally urgent topic but is marred by fatal methodological flaws. The use of "simulated interviews" without clarification or validation, combined with unverified data sources and inadequate statistical rigor, renders the findings unreliable. While the ethical framing around epistemic justice is commendable, it cannot compensate for foundational weaknesses.

\\\\ \\ **Recommendation: Reject** **Justification**: - The study fails to meet basic standards of scientific validity and reproducibility. - Ethical concerns (e.g., unresolved IRB status, potential misrepresentation of data) are unacceptable for a high-impact journal. - The incremental novelty does not outweigh these deficits.

If the authors address the major issues above, a resubmission after rigorous revision and external validation could be considered. However, in its current form, the manuscript is not suitable for publication.

—

Reviewer 01 *Expert Peer Reviewer for [Journal Name]* *Confidential Comments to Editor*: This manuscript requires substantial ethical and methodological overhaul before it can be taken seriously. The political sensitivity of the topic demands even greater rigor to avoid perceived bias.

AI Reviewer 2 - Critical Review

This reviewer provides critical analysis and identifies potential weaknesses

****Overall Impression****

My immediate reaction is one of significant methodological concern. While the topic of epistemic justice in humanitarian data is timely and important, this manuscript suffers from fundamental flaws in research design and execution that undermine its scientific validity. The study presents as a mixed-methods investigation but relies on "simulated interviews" - a methodological red flag that raises serious questions about data authenticity. The quantitative analysis, while technically sound, examines a dataset with questionable representativeness and appears disconnected from the stated research questions about trustworthiness. This feels like an overhyped weak study attempting to leverage a politically charged context to mask methodological deficiencies.

****Technical \ Scientific Assessment****

****A. Problem Definition (Score: 3/5)**** The research question is clearly motivated by important theoretical frameworks (epistemic justice, communicative competence) and addresses a genuine problem in humanitarian practice. However, the connection between the quantitative analysis of damage data and the qualitative investigation of trustworthiness remains conceptually weak. The authors fail to convincingly argue why statistical correlations between damage variables should inform our understanding of institutional trust.

****B. Methodological Soundness (Score: 1/5)**** The use of "simulated interviews" represents a fatal methodological flaw. Without transparency about what "simulated" means (AI-generated? role-play? fabricated narratives?), the qualitative data cannot be considered valid evidence. The sampling strategy (n=18) lacks justification for sufficiency, and the absence of details about interview simulation procedures violates basic research transparency standards. The quantitative methods are technically adequate but applied to data that may suffer from selection bias and unknown verification protocols.

****C. Results \ Evidence (Score: 2/5)**** The quantitative results demonstrate statistical coherence but don't actually address the core research questions about trustworthiness. High correlations between damage variables could simply reflect coordinated reporting rather than data quality. The qualitative "findings" are fundamentally compromised by the simulation methodology. Claims about "moral-trust statements" and their statistical relationship to correlation coefficients are speculative at best, given the artificial nature of the interview data.

****D. Contribution to the Field (Score: 2/5)**** While the integration of epistemic justice frameworks with humanitarian data science is potentially valuable, the methodological flaws prevent this study from making a meaningful contribution. The central finding - that community data shows internal consistency - is unsurprising and doesn't substantially advance our understanding of trust establishment.

****E. Writing \ Presentation (Score: 3/5)**** The paper is generally well-written and logically organized, with appropriate theoretical framing. However, the persistent use of "simulated interviews" without adequate explanation creates confusion and undermines credibility.

****F. Ethical \ Transparency Standards (Score: 1/5)**** The failure to disclose the nature of "simu-

lated interviews” represents a serious transparency violation. There’s no indication of whether IRB approval was obtained for human subjects research involving simulated participants. The dataset is cited but without critical examination of its collection methodology or potential biases.

****Strengths****

- Important theoretical framing combining epistemic justice with humanitarian data practices - Clear articulation of the problem of institutional skepticism toward community data - Technically competent quantitative analysis - Appropriate use of mixed-methods terminology and design description

****Weaknesses****

****Major Flaws:**** - Unvalidated ”simulated interview” methodology that compromises all qualitative findings - Disconnect between quantitative results and research questions about trustworthiness - Lack of critical examination of the Kaggle dataset’s collection methodology and potential biases - Insufficient justification for sample size and sampling strategy - Overstated claims about the relationship between statistical correlations and trust establishment

****Minor Flaws:**** - Inconsistent citation formatting in references - Vague description of ”NVivo-style coding” without specific procedures - Unclear how ”moral-trust statements” were operationalized and quantified

****Recommendations for Improvement****

1. ****Replace or substantially validate the qualitative component:**** Conduct genuine interviews with actual participants or provide exhaustive documentation of simulation procedures with validation evidence.
2. ****Strengthen theoretical connections:**** Develop clearer conceptual links between statistical coherence and trustworthiness, perhaps through mediation analysis or additional validation metrics.
3. ****Address dataset limitations:**** Critically examine the Kaggle dataset’s collection methodology, potential biases, and representativeness.
4. ****Improve methodological transparency:**** Provide detailed protocols for all data collection and analysis procedures, including explicit ethical oversight documentation.
5. ****Converge methods more meaningfully:**** Design integration procedures that actually test hypotheses about the relationship between data quality and trust establishment.

****Verdict****

****Overall Score: 1/5 - Strong Reject****

****Justification:**** This paper cannot be accepted in its current form due to the fundamental methodological flaw of using ”simulated interviews” without adequate explanation or validation. This compromises the entire qualitative component and, by extension, the mixed-methods integration. Even if this issue were resolved, significant concerns remain about the conceptual disconnect between the quantitative findings and the research questions about trustworthiness. The study attempts to address an important problem but fails to provide scientifically valid evidence for its conclusions. A complete redesign with authentic data collection would be necessary for reconsid-

eration.

Human Reviewer

Expert human review providing domain-specific insights

Reviewer: Dr. Thames

Scores (out of 10): - Question 1: 0.5 - Question 2: 0 - Question 3: 0 - Question 4: 0.25 - Question 5: 0 - Question 6: 0 - Question 7: 0

Total Score: 0.75 / 10

Reviewer Notes:

Submitted at: 2025-12-25 01:28:48

Aggregate Review

Combined assessment from all reviewers

Aggregate Review Report

This report combines the assessments from all three reviewers:

Individual Scores: - AI Reviewer 1 (Comprehensive): 6.0 / 10 - AI Reviewer 2 (Critical): 0.0 / 10
- Human Reviewer (Dr. Thames): 0.75 / 10

Aggregate Score: 2.25 / 10

Overall Assessment: The paper received an average score of 2.25 out of 10 across all three reviewers.

Final Decision: Reject

Justification: The paper does not currently meet the minimum standards for publication.

Generated: 2026-01-06 14:16:38