

REVIEWER'S REPORT**Manuscript No.: JNHST-047****Title : Generative AI and the Social Sciences: Rethinking Methodology with Large Language Models and Synthetic Populations****Recommendation:**

Accept as it is

Accept after minor revision.....

Accept after major revision

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

Reviewer's ID: JNJST-04Detailed Reviewer's Report**Overall Evaluation**

This manuscript addresses a highly contemporary and increasingly important methodological question: how large language models (LLMs) and synthetic populations may reshape social science research. The paper examines the epistemological foundations of LLM-based social simulation, reviews emerging applications in survey research, multi-agent modeling, and social digital twins, and discusses challenges related to probabilistic inference, representativeness, diversity collapse, and population-level validity. The manuscript ultimately argues for a pragmatic and cautious methodological framework in which LLMs are treated as powerful pattern-matching systems that complement, rather than replace, traditional social scientific inference.

The topic is both timely and relevant. As generative AI becomes increasingly integrated into research workflows, social scientists face important questions regarding validity, reliability, representativeness, and the epistemological status of AI-generated data. The manuscript demonstrates awareness of recent developments in computational social science and successfully brings together literature from machine learning, social simulation, survey methodology, and philosophy of social science. Its central message—

REVIEWER'S REPORT

that enthusiasm for synthetic populations should be balanced with methodological caution—is sensible and likely to resonate with a broad interdisciplinary audience.

A notable strength of the paper is its attempt to move beyond purely technical discussions and engage directly with questions of inference and scientific validity. The discussion of probabilistic inference, diversity collapse, population-level realism, and evaluation frameworks reflects a commendable effort to connect emerging AI technologies with long-standing methodological concerns in social research. The figures and conceptual diagrams further help organize the discussion and make complex issues accessible to readers.

At the same time, the manuscript functions primarily as a conceptual review and synthesis rather than an original empirical contribution. While this is not necessarily a weakness, several arguments would benefit from deeper critical engagement and stronger analytical development. In some sections, the paper reads as a summary of recent studies rather than a sustained scholarly intervention. There are also a number of language, stylistic, and structural issues that require attention before publication. Nevertheless, the manuscript addresses an important emerging field and offers a useful contribution to ongoing methodological debates.

Strengths of the Paper

The most significant strength of the manuscript is the relevance of its subject matter. Questions concerning the use of generative AI in social science are rapidly becoming central to contemporary methodological discussions. By focusing on synthetic populations, survey simulation, and AI-assisted social modeling, the paper engages with issues that are likely to shape future research practices across multiple disciplines.

The manuscript also succeeds in presenting a balanced perspective. Rather than adopting either an overly optimistic or overly skeptical position, the author acknowledges both the opportunities and limitations associated with LLM-based social research. This balanced approach strengthens the credibility of the paper and makes its conclusions more persuasive.

Another strength lies in the breadth of literature covered. The paper synthesizes a wide range of recent scholarship on LLMs, synthetic survey generation, multi-agent systems, digital twinning, population realism, and evaluation frameworks. This interdisciplinary engagement allows readers to gain a broad overview of a rapidly evolving field.

REVIEWER'S REPORT

The discussion of epistemological challenges is particularly valuable. The treatment of probabilistic inference, order dependence, representational bias, and diversity collapse demonstrates an awareness that methodological questions cannot be reduced to technical performance metrics alone. The manuscript appropriately emphasizes that social scientific validity requires more than predictive accuracy.

The proposed pragmatic framework and methodological guardrails provide a useful concluding contribution. The recommendations regarding validation, subgroup calibration, diversity auditing, and hybrid research designs offer practical guidance that researchers may find valuable when incorporating generative AI into empirical work.

Areas for Improvement

Although the manuscript is presented as a methodological contribution, its precise contribution could be articulated more clearly. At times it is difficult to determine whether the paper aims to provide a literature review, a conceptual framework, a methodological critique, or a normative guideline for future research. Clarifying the primary objective of the paper early in the introduction would strengthen its overall coherence.

The paper would also benefit from deeper critical engagement with some of the literature it cites. Many studies are introduced and summarized, but relatively little attention is devoted to comparing competing perspectives or evaluating the strength of their evidence. In several places, findings from recent studies are accepted at face value without sufficient discussion of methodological limitations, conflicting results, or unresolved debates within the field. A more critical synthesis would enhance the scholarly depth of the manuscript.

A related issue concerns the extensive reliance on very recent and, in several cases, forthcoming or preprint sources. While this is understandable given the rapidly evolving nature of generative AI research, the manuscript should acknowledge more explicitly the provisional status of some claims and findings. The field remains highly dynamic, and conclusions drawn from emerging studies may require careful qualification.

The discussion of statistical inference could be developed further. The paper correctly distinguishes pattern matching from probabilistic inference, but some of the methodological implications remain underexplored. Readers would benefit from a more detailed explanation of how synthetic data generation intersects with established concerns in sampling theory, measurement validity, causal inference, and

REVIEWER'S REPORT

external validity. Expanding these connections would strengthen the contribution to social science methodology.

While the manuscript discusses representational bias extensively, the treatment of ethical considerations remains relatively limited. Questions concerning accountability, transparency, consent, algorithmic fairness, and the potential misuse of synthetic populations deserve greater attention, especially given the policy implications of social simulation and digital twinning. A dedicated discussion of ethical implications would enhance the completeness of the analysis.

The organization of some sections could also be improved. Certain ideas appear repeatedly across the paper, particularly discussions concerning diversity collapse, demographic bias, and population-level validity. Consolidating these discussions would reduce redundancy and improve the overall flow of the manuscript.

Finally, the manuscript would benefit from careful language editing. There are numerous grammatical inconsistencies, punctuation problems, spacing errors, duplicated phrases, and awkward sentence constructions throughout the text. For example, some sections contain repeated sentences, inconsistent terminology regarding LLMs, and formatting irregularities in tables and figures. These issues do not undermine the substantive argument but do affect readability and presentation.

Final Recommendation**Accept with Minor Revisions**

This manuscript addresses an important and rapidly developing area of social science methodology. It provides a useful synthesis of current debates surrounding large language models, synthetic populations, and computational social simulation while maintaining an appropriately cautious perspective on methodological validity. The discussion is timely, interdisciplinary, and conceptually engaging. Although the paper would benefit from greater analytical depth, clearer articulation of its contribution, stronger critical engagement with the literature, and careful language revision, these concerns are largely refinements rather than fundamental shortcomings. With minor revisions, the manuscript has the potential to make a valuable contribution to discussions on the responsible integration of generative AI into social science research.