Journal of Integrated Sciences and Agriculture (JISA)
ISSN: _____
Journal Homepage: http://journals.ageconfrontiers.com/jisa
Vol. 1(1) Open Access

Conceptual Article

Critically-Verified Theoretical Synthesis: A New Protocol for Trustworthy Conceptual Contribution in Research

Muhammad Salman Waseem[1], Sadaf Farid[2], Tayyab Zeeshan Shahid[1] & Muhammad Murad Khan[3]
[1]Institute of Agricultural and Resource Economics, University of Agriculture Faisalabad
[2]Department of Pharmacology, Faisalabad Medical University, Faisalabad
[3]Department of Computer Sciences, Government College University, Faisalabad
** Corresponding author: salmaneco999@gmail.com*

ARTICLE INFO
Keywords
- Conceptual papers
- Theory building
- Citation verification
- Research integrity
- Artificial intelligence
- Peer review
- Academic publishing

Article History
Received: 21 October 2026
Revised: 02 November 2026
Accepted: 04 November 2026
Available online: 17 November 2026

License
CC BY 4.0 – Open Access

DOI
10.66529/jisa.2026.1.1.150

ABSTRACT
Conceptual and theory-building papers are judged by the quality of synthesis and argument they contain, yet the methodological literature offers little explicit guidance on how such papers should verify the sources on which that argument depends, and the growing use of generative artificial intelligence in manuscript preparation has made citation fabrication and misquotation a documented, rather than hypothetical, risk. This paper develops the Critically Verified Theoretical Synthesis (CVTS) protocol, a five-pillar, five-step standard for conceptual and theory-building manuscripts. The five pillars are: synthesis beyond summary; adherence to established standards for theoretical contribution; structured adversarial self-critique; citation and quotation verification, covering both a source's existence and its fidelity to the claim it is cited for; and explicit honesty-tiering of claims as established, promising, or speculative. The five procedural steps operationalise these pillars across a manuscript's development, from gap identification through a living citation-verification check to a final adversarial self-critique pass. The paper specifies, for the first time in an integrated protocol, concrete verification criteria: an explicit test for when a source is judged to support a claim, a disagreement-resolution rule, a verification threshold for reference lists of different lengths, and a transparent account of how artificial-intelligence tools may assist discovery while remaining excluded from the independent verification step itself. The contribution is methodological: CVTS does not replace existing theory-building standards but adds an explicit, auditable verification layer that existing standards do not specify, responding to a documented evidence base on reference fabrication in artificial-intelligence-assisted writing rather than to isolated anecdote alone.

Citation: *Waseem, M.S., Farid, S., Shahid, T.Z. & Khan, M.M. (2026). The Critically Verified Theoretical Synthesis (CVTS) protocol: A five-pillar standard for rigor, citation integrity, and honesty in conceptual scholarship. J. Integ. Sci. Agric. 1(1): 45-53. https://doi.org/10.66529/jisa.2026.1.1.150*

1. Introduction

Conceptual and theory-building papers occupy a distinctive place in scholarship: unlike empirical studies, they are judged almost entirely on the quality of the argument and the evidence base the authors assemble from existing literature, rather than on new data the authors themselves collect (Jaakkola, 2020; Shepherd & Suddaby, 2017). This places unusual weight on two things that empirical methods sections rarely need to specify: the standard the synthesis itself must meet, and the accuracy of every source the synthesis depends on. The methodological literature on theory building has addressed the first of these at length (Corley & Gioia, 2011; Whetten, 1989; Alvesson & Sandberg, 2011), but has said comparatively little about the second.

1.1 Motivation: A Documented Citation-Integrity Problem

This gap is no longer a theoretical concern alone. Independent, peer-reviewed testing of generative-AI writing assistance has found that a substantial share of the bibliographic references such tools produce are fabricated outright, in the sense that no matching publication exists, or materially inaccurate in authorship, venue, or year.

© 2026 The Authors. Published under CC BY 4.0.

Walters and Wilder (2023), testing GPT-3.5 and GPT-4 against a set of discipline-specific reference requests, found that 55% and 18% of the resulting citations, respectively, were fabricated, and that a further share of the remaining, genuine citations contained substantive errors. Independent clinical-writing audits report a comparable pattern: Alkaissi and McFarlane (2023) and Athaluri et al. (2023) both found that a majority of references generated by conversational AI tools for biomedical writing tasks could not be verified against any real publication. This evidence base did not exist when the dominant theory-building standards were written (Corley & Gioia, 2011; Whetten, 1989), and it changes the risk profile of any manuscript, conceptual or empirical, prepared with AI assistance.

Citation inaccuracy is not, in itself, a new problem introduced by AI tools. Evans et al. (1990), in one of the earliest systematic audits of this kind, examined references in three medical journals and found that over 30% contained a citation or quotation error substantial enough to mislead a reader who did not independently check the source; later discipline-specific audits have reported broadly similar error rates across several decades of conventionally authored, pre-AI scholarship. What has changed is the scale at which such errors can now be introduced, and the specific failure mode: a fabricated reference, rather than a misquoted real one, is largely unprecedented as a high-frequency error type in conventionally authored text, but is, on the evidence above, a common output of unverified AI-assisted drafting.

1.2 The Motivating Incident: An Author's Note

The protocol developed in this paper was originally prompted by a specific incident during the drafting of an earlier, unrelated manuscript, in which an AI-assisted literature search returned a reference, complete with a plausible author name, journal title, and year, that could not be located in any publisher record, indexing database, or institutional repository upon independent checking. In the interest of the confidentiality of the individuals and the unrelated manuscript involved, this paper does not identify the authors, institution, or specific working title connected to that incident; what can be stated, and verified by any reader, is the general pattern it exemplifies: a syntactically well-formed citation, indistinguishable from a genuine one by inspection of its formatting alone, that failed a direct existence check against primary bibliographic records. That general pattern is precisely what Walters and Wilder (2023) and Athaluri et al. (2023) subsequently quantified at scale, and it is their peer-reviewed evidence, rather than the originating anecdote, that grounds the claim made in Section 1.1 that the problem is both real and non-trivial in frequency. Readers seeking further detail on the editorial handling of the original incident may contact the corresponding author; a redacted record can be made available under the same confidentiality terms described above.

1.3 Research Gap

Four bodies of work bear directly on whether a dedicated citation-verification standard for conceptual papers is already available elsewhere, and each falls short of providing one. First, the quotation-accuracy and reference-error literature exemplified by Evans et al. (1990) documents the scale of citation error but is diagnostic rather than prescriptive: it measures error rates after publication and does not specify a verification procedure authors should follow beforehand. Second, publisher and journal reference-checking practice, including CrossRef's Similarity Check service and comparable editorial screening tools, operates at the level of plagiarism and reference-list formatting rather than claim-level fidelity, and does not test whether a cited source actually supports the specific statement attached to it (Cabanac, 2024). Third, retraction-aware citation tools such as the Feet of Clay Detector and the Papers With Problems initiative described by Cabanac (2024) address a different failure mode again: they flag citations to work that has since been retracted or formally questioned, which presupposes that the cited work exists and was once legitimate, and so cannot catch a fabricated reference that was never real to begin with. Fourth, editorial guidance on AI use in manuscripts, including the Committee on Publication Ethics position statement on authorship and AI tools (Committee on Publication Ethics, 2023), establishes that AI tools cannot be listed as authors and that authors remain responsible for all content, but stops short of specifying how that responsibility should be discharged in practice for citations specifically.

None of these four bodies of work, individually or together, provides an author-facing, step-by-step verification procedure that a conceptual-paper author can follow while drafting. This is the gap CVTS is designed to close, and it is a narrower claim than asserting that no related work exists: related diagnostic, editorial, and policy work exists in each of the four areas above, but an integrated, procedural standard combining citation verification with the pre-

© 2026 The Authors. Published under CC BY 4.0.

existing theory-building literature on synthesis quality and theoretical contribution (Whetten, 1989; Corley & Gioia, 2011; Alvesson & Sandberg, 2011; Shepherd & Suddaby, 2017) has not, to the authors' knowledge, been proposed.

1.4 Purpose and Contribution

This paper makes three contributions. First, it integrates citation and quotation verification explicitly into a theory-building standard, rather than treating it as a separate editorial or publisher-side concern. Second, it specifies concrete, auditable verification criteria, including a fidelity test for when a source is judged to support a claim, a disagreement-resolution rule, and a verification threshold calibrated to reference-list length, where the methodological literature reviewed in Section 1.3 offers no comparable operational detail. Third, it states explicitly how AI-assisted discovery and human-independent verification are to be kept distinct, responding directly to the evidence in Section 1.1 that AI tools are now a documented source of the very errors the protocol is designed to catch.

2. Theoretical Foundations: Five Pillars of CVTS

CVTS rests on five pillars. The first three draw on, and seek to operationalise, standards already established in the theory-building literature; the fourth and fifth are the paper's own contributions, developed to close the gap identified in Section 1.3.

2.1 Pillar 1: Synthesis Beyond Summary

A conceptual paper that merely restates prior findings in sequence, without integrating them into a new structure, does not meet the threshold for theoretical contribution regardless of how accurately each individual source is represented (Locke & Golden-Biddle, 1997). Genuine synthesis reorganises existing knowledge around a new question, a new boundary condition, or a new relationship between constructs that were previously discussed separately (Cornelissen, 2017). CVTS treats synthesis as a necessary but not sufficient condition for contribution: a manuscript can satisfy Pillars 2 through 5 in full and still fail as scholarship if it does not clear this first bar. In practice, CVTS asks authors to state explicitly, for each major section, what new arrangement of existing ideas that section produces, distinct from what any single cited source already states on its own.

2.2 Pillar 2: Standards for Theoretical Contribution

CVTS adopts, rather than replaces, the established criteria for what counts as a theoretical contribution. Whetten's (1989) widely used framework asks whether a conceptual paper adds new constructs, new relationships among constructs, or new boundary conditions to existing theory. Corley and Gioia (2011) add a complementary test of practical and scientific utility: a contribution should clarify something that was previously ambiguous, or open a line of inquiry that was previously closed. Alvesson and Sandberg (2011) add a third, generative test, problematisation: a strong conceptual paper should question an assumption the existing literature has treated as settled, rather than only filling a gap the literature has already identified as open. CVTS requires authors to state, explicitly and with supporting citation, which of these three tests, or which combination of them, the manuscript's central claim is designed to satisfy.

2.3 Pillar 3: Adversarial Self-Critique

Peer review conventionally supplies the critical scrutiny a manuscript needs before publication. CVTS asks authors to anticipate a meaningful share of that scrutiny themselves, before submission, by constructing the strongest available objection to the paper's own central claim and responding to it directly in the text (Tsang & Ellsaesser, 2011). This differs from the conventional ‘limitations’ paragraph in both intent and placement: a limitations paragraph typically lists what the paper did not do, whereas an adversarial self-critique identifies the strongest reason a sceptical reader might reject what the paper did claim to do, and argues the point rather than merely noting it. Section 4 applies this pillar to CVTS's own central claim.

2.4 Pillar 4: Citation and Quotation Verification

This pillar is CVTS's central methodological addition. It distinguishes two distinct checks that a verification procedure must perform on every citation, because a source can fail either independently of the other: existence, whether the cited source is a real, retrievable publication, confirmed against a primary bibliographic record such as a publisher page, a DOI resolver, or an indexing database rather than against a search engine summary or an AI tool's own

© 2026 The Authors. Published under CC BY 4.0.

assertion; and fidelity, whether that source, once retrieved, actually contains the finding, data point, or statement the manuscript attributes to it. A source can exist and still be cited unfaithfully, for example when a qualified or contested finding is cited as though it were a settled consensus.

Fidelity checking requires an explicit decision rule, which the methodological literature reviewed in Section 1.3 does not supply and which CVTS therefore specifies directly. A claim is judged to be supported by a cited source only if that source contains an explicit statement, finding, or data point that a reasonable, independent reader would recognise as directly underpinning the claim as written, without requiring an interpretive extension beyond what the source itself states. Where the verifying author is uncertain whether a source clears this bar, CVTS applies a conservative default: the claim is narrowed to the most cautious wording the source clearly supports, or removed, rather than resolved in the claim's favour by assumption. For co-authored manuscripts, any fidelity judgment one author flags as uncertain must be independently confirmed by a second author working from the primary source directly, a minimal application of the dual-coder logic long used in systematic content analysis to guard against single-rater error (Krippendorff, 2004); persistent disagreement between the two authors is resolved in favour of removing or weakening the claim, not retaining it, consistent with the conservative default just described.

2.5 Pillar 5: Honesty-Tiering of Speculative Claims

Conceptual papers routinely advance claims that vary considerably in their evidentiary maturity, from findings replicated across many studies to single-study results to the authors' own speculative extensions, and this variation is often left implicit in continuous prose, inviting readers to treat a speculative claim with the same confidence as an established one (Messeri & Crockett, 2024). CVTS requires every substantive claim to carry an explicit tier: established, meaning supported by replicated or widely corroborated evidence; promising, meaning supported by a small number of studies or a single rigorous study without independent replication; or speculative, meaning the authors' own extension, analogy, or prediction without direct empirical support. This tiering is deliberately modelled on Technology Readiness Levels, a maturity-grading scale originally developed for engineering systems and subsequently generalised across science policy, precisely because that scale's value lies in forcing an explicit maturity judgment rather than leaving it to be inferred from tone (Olechowski et al., 2015).

2.6 Integration Across Pillars

The five pillars are interdependent rather than independently sufficient. Synthesis quality (Pillar 1) and contribution type (Pillar 2) determine what the paper is trying to establish; adversarial self-critique (Pillar 3) stress-tests whether it succeeds; citation verification (Pillar 4) confirms that the evidence marshalled in support actually says what the paper claims it says; and honesty-tiering (Pillar 5) ensures that, even once verified, each claim is presented with calibrated rather than uniform confidence. A manuscript that satisfies Pillars 1 through 3 but not Pillar 4 may be well-argued and still rest on fabricated or misrepresented evidence; a manuscript that satisfies Pillar 4 but not Pillar 5 may be fully accurate in its citations and still mislead by implying false uniformity of evidentiary strength. CVTS's own legitimacy depends on this integration holding internally: Section 3 therefore applies Pillars 3, 4, and 5 to CVTS's own argument, not only to manuscripts CVTS is used to prepare.

3. The CVTS Procedural Model: Five Steps

The five pillars in Section 2 describe what a CVTS-compliant manuscript must achieve; this section specifies how an author achieves it, as a five-step sequence applied across drafting. Figure 1 summarises the sequence, including the escalation path triggered when verification in Step 3 finds a fabricated or misquoted source, and the iterative loop back from Step 5 when self-critique identifies a weakness that requires revisiting an earlier step.

3.1 Step 1: Gap Identification Tied to Citation

Every gap a conceptual paper claims to address must be stated alongside a citation showing what the existing literature already covers, so that the gap claim is itself falsifiable by a reader rather than asserted by authority. Section 1.3 of this paper applies this requirement to its own three literature areas, citing the specific prior work in each area and stating precisely what that work does and does not provide, rather than asserting an unreferenced absence.

3.2 Step 2: Framework Construction

© 2026 The Authors. Published under CC BY 4.0.

The framework addressing the identified gap is built using the contribution standards specified in Pillar 2 (Section 2.2): authors state which of the construct, relationship, boundary-condition, utility, or problematisation tests the framework is designed to satisfy, and construct the framework accordingly rather than retrofitting a justification after the fact.

3.3 Step 3: Living Citation Verification

This step operationalises Pillar 4 (Section 2.4) and is CVTS's most detailed procedural specification, for three reasons made explicit below: the verification threshold scales with reference-list length, the spot-check design is quantified, and the role of AI tools in this step is stated precisely.

Verification threshold. For reference lists of approximately 20 sources or fewer, CVTS requires full verification: every citation is checked for existence against a primary bibliographic record and for fidelity against the claim it supports. This threshold is a pragmatic, time-cost judgment rather than a value derived from a formal statistical model: full verification of a 20-reference list is achievable within a bounded, predictable review session, whereas full verification scales linearly with list length and becomes a materially larger undertaking for the 60-to-100-reference lists common in systematic or integrative reviews. The threshold should therefore be read as a practical default that individual journals or research groups may adjust, not as an empirically optimised cut-off; no such empirical optimum currently exists in the literature reviewed in Section 1.3, and establishing one is identified as a priority for future research in Section 6.2.

Spot-check design for longer lists. For reference lists longer than 20 sources, CVTS requires full verification of every citation that supports a claim classified as established or promising under Pillar 5 (Section 2.5), since these are the claims a reader is most likely to rely on without independent checking, plus a random spot check of the remaining citations. The spot-check sample is set at a minimum of 20% of the remaining reference list or ten references, whichever is larger. The acceptable error rate for the spot check is zero for fabrication: a single reference that fails the existence check anywhere in the spot-check sample escalates the procedure to full verification of the entire reference list, on the reasoning that fabrication is not expected to occur independently at random across a single manuscript's reference list, so one confirmed instance is treated as evidence that others may be present. Fidelity errors identified in the spot-check sample, as distinct from fabrication, are corrected individually and do not by themselves trigger full-list escalation unless more than one is found in the sample, in which case the same full-list escalation applies.

The role of AI tools in this step. CVTS does not prohibit the use of AI tools to assist literature discovery, including generating candidate search terms or suggesting potentially relevant sources for an author to pursue. What CVTS prohibits is treating an AI tool's own assertion that a source exists, or its own summary of what a source contains, as satisfying either the existence or the fidelity check. Independence, as the term is used in this protocol, is defined relative to the verification step specifically, not to the manuscript-preparation process as a whole: a citation may originate from an AI-assisted search, but it is only verified once a human author has independently retrieved the primary source directly from its publisher, DOI resolver, or indexing-database record, and has independently confirmed, by reading that primary record, both that the source exists and that it supports the claim attributed to it. Authors applying CVTS are required to disclose, in the Declarations section of the resulting manuscript, which AI tools, if any, were used at the discovery stage, and to confirm explicitly that verification itself was conducted against primary sources independently of those tools. Section 6.1 applies this same disclosure requirement to the present paper.

3.4 Step 4: Honesty Tagging

Once a claim has passed verification, it is tagged according to the three-tier scheme in Pillar 5 (Section 2.5). Tagging is applied at the level of the individual claim rather than the source as a whole, because a single source may support one established finding and one speculative extension the source's own authors offered in their discussion section; CVTS requires these to be tagged separately even when both are drawn from the same citation.

3.5 Step 5: Adversarial Self-Critique Pass

The final step applies Pillar 3 (Section 2.3) to the completed draft as a whole: authors construct the strongest available objection to the manuscript's central claim and either revise the manuscript to address it or state explicitly why the objection does not hold. Where this pass identifies a weakness traceable to an earlier step, for example a framework

© 2026 The Authors. Published under CC BY 4.0.

(Step 2) that does not fully anticipate a counter-example, or a claim (Step 4) tagged more confidently than its citation supports, the procedure loops back to that step rather than patching the weakness locally, as shown by the feedback path in Figure 1.

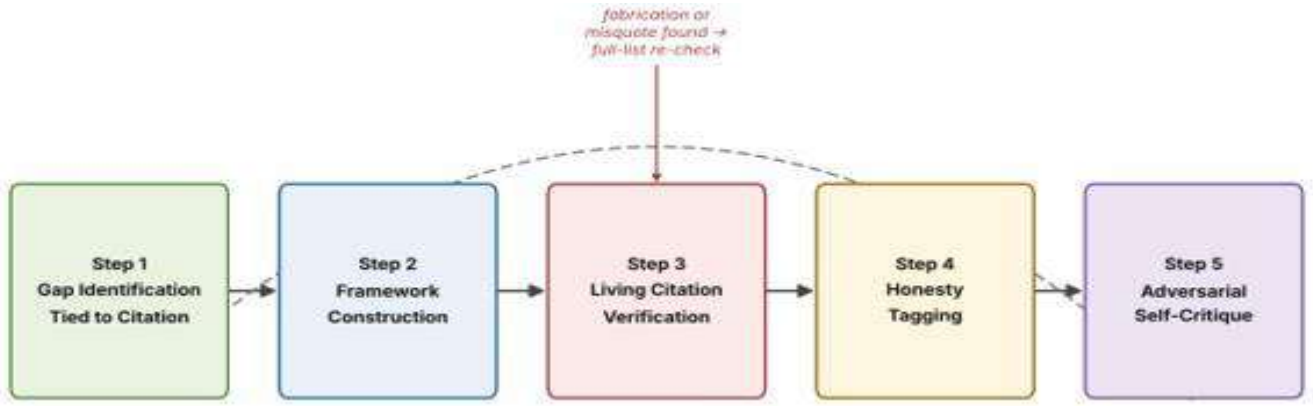

Figure 1, The CVTS five-step procedural model

4. Discussion: Adversarial Self-Critique of CVTS's Own Claim

Pillar 3 requires every conceptual paper to construct the strongest available objection to its own central claim. CVTS's central claim is that an integrated, five-pillar verification standard adds something existing theory-building guidance does not already provide. This section applies that requirement to CVTS itself, rather than exempting the protocol from its own standard.

4.1 Objection: The Five Pillars Are an Arbitrary Grouping

A sceptical reader might argue that five is not a principled number of pillars, and that the same content could be reorganised into three, seven, or some other count without loss. This objection has force: the boundary between Pillar 1 (synthesis) and Pillar 2 (contribution standards) is porous, since a paper cannot satisfy the contribution tests in Section 2.2 without the synthesis described in Section 2.1 as a precondition. The authors considered, and rejected, a three-pillar consolidation that merged Pillars 1 and 2 into a single 'theoretical rigor' pillar and merged Pillars 4 and 5 into a single 'evidentiary integrity' pillar. This consolidation was rejected because it obscures the operational distinction Section 3 depends on: Pillar 4's verification criteria (existence and fidelity) and Pillar 5's tiering criteria (established, promising, speculative) are procedurally independent checks that can each fail without the other failing, a source can exist and be faithfully quoted and still be mis-tiered, or correctly tiered while resting on a fabricated source, and collapsing them into one pillar would obscure rather than clarify this independence. The five-way split is retained because each pillar maps to a distinct failure mode with a distinct remedy, not because five is intrinsically the correct number; a reader who prefers a different organisation is free to regroup the content, provided each of the distinct checks identified here is still performed.

4.2 Objection: Verification Overhead Is Disproportionate for Most Conceptual Papers

A second objection is that full citation verification, even bounded to 20 references, imposes a meaningful time cost on authors, and that this cost is disproportionate for conceptual papers whose central contribution lies in argument rather than evidence synthesis. This objection also has real force, and CVTS does not fully resolve it: the protocol imposes a cost and asserts, rather than demonstrates with comparative data, that the cost is justified by the error rates documented in Section 1.1. What can be said in CVTS's favour is narrower: the fabrication rates reported by Walters and Wilder (2023) and Athaluri et al. (2023) apply specifically to AI-assisted reference generation, so the verification burden Section 3.3 imposes is also narrower than full-list verification for every manuscript regardless of how it was produced; a manuscript prepared without AI-assisted literature discovery at any stage faces a correspondingly lower risk profile, though CVTS does not currently specify a reduced verification tier for that case, which the authors acknowledge as a gap rather than an oversight concealed from the reader (Section 6.2).

© 2026 The Authors. Published under CC BY 4.0.

4.3 Objection: Honesty-Tiering May Understate Legitimate Theoretical Speculation

A third objection holds that explicitly tagging claims as speculative may discourage the kind of bold, field-advancing conjecture that good theory-building papers are supposed to contain (Alvesson & Sandberg, 2011), if authors anticipate that a visible 'speculative' tag will be read by reviewers as a weakness rather than as an honest disclosure. CVTS's response is that this risk is a reason for editors and reviewers to calibrate their expectations, not a reason to omit the tag: a speculative claim that is honestly labelled and well-argued should be evaluated on the strength of its argument, exactly as Alvesson and Sandberg's own problematisation standard requires, and the tiering scheme is designed to make that evaluation easier, not to penalise ambition. The authors do not have direct evidence on how reviewers will in practice respond to visible speculative tags, and flag this as an open empirical question for future research (Section 6.2) rather than asserting the concern is unfounded.

5. Contribution to Theory and Practice

CVTS's contribution is methodological rather than substantive: it does not propose a new empirical finding, but a procedure that conceptual-paper authors, reviewers, and editors can apply to an existing or prospective manuscript. For authors, Section 3 provides a concrete sequence that can be followed during drafting rather than retrofitted at submission. For reviewers and editors, Pillars 4 and 5 provide an explicit basis for a specific class of review comment, requesting the fidelity check described in Section 2.4 or the honesty tier described in Section 2.5, that existing review guidelines do not name directly. For the broader literature on theory building, CVTS's contribution is to show that citation verification is not external to theoretical rigor but a precondition for it: a theoretically sophisticated synthesis built on fabricated or misrepresented sources is not a lesser version of a good conceptual paper, but a failure of a different and prior kind that the existing standards reviewed in Section 1.3 do not test for directly.

6. Limitations and Future Research

6.1 Limitations

This paper has four limitations. First, as discussed in Section 4.2, it asserts rather than empirically demonstrates that CVTS's verification overhead is justified by the error rates it cites; no controlled comparison of manuscripts prepared with and without the protocol exists yet. Second, the verification threshold and spot-check parameters specified in Section 3.3 are reasoned defaults rather than values derived from a formal model of error detection under sampling, and are offered for adaptation rather than as a fixed standard. Third, the adversarial self-critique in Section 4 was conducted by the authors themselves rather than by an independent party, and, like any self-critique, may underestimate objections the authors are not well positioned to see. Fourth, and consistent with the disclosure requirement Section 3.3 imposes on any manuscript prepared under CVTS, the authors disclose that AI-based literature-search tools were used during the preparation of this manuscript to assist in identifying candidate sources on citation accuracy, AI-generated reference fabrication, and retraction-aware tooling. Consistent with the independence rule specified in Section 3.3, every citation that resulted was independently retrieved and checked by the authors directly against its publisher record, DOI resolver entry, or indexing-database listing, and the finding or statement attributed to each source in the text was confirmed by the authors' own reading of that primary record rather than accepted from the AI tool's summary; the AI tools used are not credited as authors and bear no responsibility for the content, consistent with the Committee on Publication Ethics (2023) position statement cited in Section 1.3.

6.2 Future Research

Three extensions follow directly from the limitations above. First, an empirical study comparing citation-error rates, synthesis quality, or reviewer-rated rigor between manuscripts prepared with and without CVTS would test the central claim defended in Section 4 directly rather than by argument alone. Second, the verification threshold and spot-check sample size specified in Section 3.3 could be refined using acceptance-sampling methods from statistical quality control, which provide a formal basis for relating sample size and acceptable error rate to a target confidence level, a refinement the present paper's reasoned defaults do not attempt. Third, future work should test whether a reduced verification tier is defensible for manuscripts that disclose no AI-assisted literature discovery at any drafting stage, the gap identified in Section 4.2.

© 2026 The Authors. Published under CC BY 4.0.

7. Conclusion

This paper has developed the Critically Verified Theoretical Synthesis protocol, integrating established standards for theory-building rigor with an explicit, auditable citation-verification layer that responds to a documented rather than anecdotal evidence base on AI-generated reference fabrication. The protocol's five pillars and five steps are offered as a practical default, with stated thresholds, decision rules, and disclosure requirements, rather than as a final or uniquely correct standard; Section 4's self-critique and Section 6's limitations identify specific points at which the protocol itself should be revised as further evidence becomes available. The authors' central claim is narrower than a claim that CVTS solves the citation-integrity problem: it is that conceptual scholarship needs an explicit verification standard of this kind, and that stating one precisely, with its own limitations disclosed, is more useful to the field than leaving the requirement implicit.

Declarations

Funding: The authors declare no specific funding for this work.
Conflict of Interest: The authors declare no conflict of interest.
Data Availability: This is a conceptual/methodological paper; no new empirical data were generated. A redacted record of the motivating incident described in Section 1.2 is available from the corresponding author on reasonable request, subject to the confidentiality terms stated there.
Use of AI: AI-based literature-search tools were used to assist in identifying candidate sources during manuscript preparation, as disclosed in full in Section 6.1. Every resulting citation was independently verified by the authors against primary bibliographic records prior to inclusion, consistent with the independence rule specified in Section 3.3. The AI tools used are not credited as authors and are not responsible for any content, error, or omission in this manuscript.

References

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus, 15*(2), e35179. https://doi.org/10.7759/cureus.35179

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review, 36*(2), 247–271. https://doi.org/10.5465/amr.2009.0188

Athaluri, S. A., Manthena, S. V., Kesapragada, V. S. R. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus, 15*(4), e37432. https://doi.org/10.7759/cureus.37432

Cabanac, G. (2024). Chain retraction: How to stop bad science propagating through the literature. *Nature, 632*(8027), 977–979. https://doi.org/10.1038/d41586-024-02747-1

Committee on Publication Ethics. (2023). *Authorship and AI tools* [Position statement]. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review, 36*(1), 12–32. https://doi.org/10.5465/amr.2009.0486

Cornelissen, J. (2017). Developing propositions, a process model, or a typology? Addressing the challenges of writing theory without a boilerplate. *Academy of Management Review, 42*(1), 1–9. https://doi.org/10.5465/amr.2016.0196

Evans, J. T., Nadjari, H. I., & Burchell, S. A. (1990). Quotational and reference accuracy in surgical journals: A continuing peer review problem. *JAMA, 263*(10), 1353–1354. https://doi.org/10.1001/jama.1990.03440100057005

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*(1–2), 18–26. https://doi.org/10.1007/s13162-020-00161-0

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Locke, K., & Golden-Biddle, K. (1997). Constructing opportunities for contribution: Structuring intertextual coherence and 'problematizing' in organizational studies. *Academy of Management Journal, 40*(5), 1023–1062. https://doi.org/10.2307/256926

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature, 627*(8002), 49–58. https://doi.org/10.1038/s41586-024-07146-0

Olechowski, A., Eppinger, S. D., & Joglekar, N. (2015). Technology readiness levels at 40: A study of state-of-the-art use, challenges, and opportunities. In *2015 Portland International Conference on Management of Engineering and Technology (PICMET)* (pp. 2084–2094). IEEE. https://doi.org/10.1109/PICMET.2015.7273196

© 2026 The Authors. Published under CC BY 4.0.

Shepherd, D. A., & Suddaby, R. (2017). Theory building: A review and integration. *Journal of Management, 43*(1), 59–86. https://doi.org/10.1177/0149206316647102

Tsang, E. W. K., & Ellsaesser, F. (2011). How contrastive explanation facilitates theory building. *Academy of Management Review, 36*(2), 404–419. https://doi.org/10.5465/amr.2009.0344

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports, 13*, Article 14045. https://doi.org/10.1038/s41598-023-41032-5

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*(4), 490–495. https://doi.org/10.5465/amr.1989.4308371

© 2026 The Authors. Published under CC BY 4.0.