Post-Publication Reference Integrity Audit

Paper: From Snapshots to Trajectories: How Agentic AI Could Redefine Student Learning Outcomes and Transform Student Success Measurement Date: 2026-03-09 Auditor: Manual WebSearch verification (Claude Code + WebSearch, independent of paper generation) Trigger: Discovery that Lin et al. (2020) was a mashup fabrication despite passing 3 rounds of integrity checking

---

Executive Summary

Metric	Count	%
Total references	68	100%
VERIFIED	36	53%
INSTITUTIONAL (partially verifiable)	11	16%
MISMATCH	17	25%
NOT_FOUND	4	6%
Total issues	21	31%

Conclusion: 3 rounds of AI-powered integrity checking (Stage 2.5, Stage 2.5 re-verification, Stage 4.5) failed to catch 21 out of 68 reference issues. This represents a 31% false-negative rate in the integrity verification system.

Why Previous Integrity Checks Failed

Root Cause Analysis

#	Design Flaw	Impact
1	Insufficient sampling (17/71 DOI checks = 24%)	54 references never had DOI verification
2	Gray-zone classification without escalation	15 references classified as “difficult to verify”and forgotten
3	Re-checks only looked at known issues	Stage 2.5 re-verification and Stage 4.5 Phase D only verified the 22 fixes
4	Context check ≠ bibliographic check	Stage 4.5 Phase B passed references that “made sense in context”without verifying bibliographic accuracy
5	No external verification	AI relied on its own “memory” instead of WebSearch, which is the source of hallucinations

The Fundamental Contradiction

Using AI to verify AI-generated citations is equivalent to having a student grade their own exam. The same training data that generates plausible-sounding fake citations also makes them seem legitimate during verification.

Detailed Findings

NOT_FOUND (4 references —likely fabricated)

#	Citation	Issue
1	AISEL (2025)	No evidence this paper exists at PACIS 2025
2	CHEA (2024)	No policy brief titled "AI and accreditation: An emerging conversation"
3	ENQA (2024)	No 2024 working paper with this title; ENQA AI publications are from 2025
4	Hou, Tsai, Hou & Chen (2020)	Book "Institutional Research in Asia-Pacific Universities"not found on Springer

MISMATCH —Wrong Authors (6 references)

#	Citation	Pattern	Issue	Correct Authors
5	Banihashem et al. (2025)	Author Mashup	"van Ginkel, Macfadyen, Savage"are not co-authors	Gasevic, Jarodzka, Joosten-ten Brinke, Drachsler
6	El-Banna et al. (2025)	Author Spoofing	"El-Banna" does not exist on this paper	Bandi, Kongari, Naguru, Pasnoor, Vilipala
7	Kestin et al. (2025)	Author Mashup	"McCarty, Callaghan, Deslauriers" are from a different Kestin paper	Klales, Milbourne, Ponti
8	Tao et al. (2026)	Author Spoofing	"Tao, Zhang, Liu, Zhao" entirely fabricated	Arunkumar V, Gangadharan G.R., Buyya R.

#	Citation	Pattern	Issue	Correct Authors
9	Lin et al. (2020)	Mashup Fabrication	Wrong initials, wrong editor, wrong book, wrong pages, wrong year	Lin A.S.R., + Chan S.J., in Hou et al. (Eds.), pp. 65-81, 2021
10	Stanford SCALE (2025)	Wrong Attribution	Not by Stanford SCALE; actual author is Lixiang Yan	Yan, L. (2025). arXiv:2508.14825

MISMATCH —Wrong Metadata (7 references)

#	Citation	Issue	Correction
11	Coates & Zlatkin-Troitschanskaia (2019)	Wrong journal, title, pages, DOI (DOI resolves to unrelated paper)	Higher Education Policy, 32, 507-512
12	Hou, Morse & Chiang (2015)	Wrong year (2012), wrong pages, wrong DOI (DOI Misdirection)	HERD 31(6), 841-857
13	IMDA (2023)	Wrong year	Published 2020, not 2023
14	Ng, D.T.K. et al. (2024)	Wrong year	Published 2021, not 2024
15	Gandara et al. (2024)	Title significantly altered	“Detecting and Mitigating”not “Examining”
16	Temper et al. (2025)	Title fabricated, DOI wrong	About regulation, not assessment transformation
17	Zhong & Zhao (2025)	Title altered, pages wrong	pp. 319-342, not 234-256

MISMATCH —Format/Title Embellishment (4 references, minor)

#	Citation	Issue
18	OpenAI & Stanford SCALE (2025)	Blog post cited as “Technical report”
19	UNESCO (2025)	Website article cited as IIEP publication
20	Sharma (2024)	Exact article title/source unconfirmed
21	Inside Higher Ed (2026)	Title is paraphrase, not exact

Hallucination Patterns Detected

Pattern	Count	References
Mashup Fabrication (PH)	4	#5, #7, #9, #4
Author Spoofing (PAC)	3	#6, #8, #10
DOI Misdirection	2	#11, #12
Temporal Masking (SH)	3	#12, #13, #14
Title Fabrication (TF)	4	#1, #2, #3, #16
Venue Exploitation	2	#11, #18

Corrective Actions Taken

1. Integrity Verification Agent v2.0 —Overhauled with:
   - Anti-Hallucination Mandate
   - Elimination of gray-zone classifications
   - Mandatory WebSearch audit trail
   - Stage 4.5 fresh independent verification
   - Known hallucination pattern library (5 types + 5 compound patterns)
2. Paper Corrections —Applied to `full_paper_zh.md` and `full_paper_zh_apa7.tex`:
   - Removed 4 NOT_FOUND references and their in-text citations
   - Corrected 6 author lists
   - Corrected 7 metadata errors
   - Corrected 2 format issues

Key Lesson

> The integrity verification system's most dangerous failure mode is not missing a completely fake reference —those are relatively easy to catch. The most dangerous failure mode is classifying a partially-matching reference as "difficult to verify"and moving on. This gray zone is where mashup fabrications thrive, because each component of the fabricated reference is individually real.

References

- Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. Scientific Reports, 13, 14045.
- GPTZero. (2026, January 21). GPTZero finds 100 new hallucinations in NeurIPS 2025 accepted papers.
- Adams, A. et al. (2026). Compound deception in elite peer review: A failure mode taxonomy. arXiv:2602.05930.