© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

Hallucinated Citations and Algorithmic Bias: Evaluating the Impact of Student AI Literacy on Epistemic Trust and Scientific Misinformation in Among Undergraduate Students in the University of Calabar, Nigeria.

DR AMOS WILLIAM OBETEN
Department: Public Administration, ORCiD:0009-0001-4139-9869

Abstract- ***The increasing integration of generative artificial intelligence (AI) into higher education has raised concerns about hallucinated citations, algorithmic bias, epistemic trust, and scientific misinformation. This study investigated student AI literacy, epistemic trust, and susceptibility to scientific misinformation among undergraduate students at the University of Calabar, Nigeria. Specifically, it examined the relationship between AI literacy and epistemic trust; the relationship between AI literacy and susceptibility to misinformation arising from hallucinated citations and algorithmic bias; the predictive influence of AI literacy, hallucinated-citation awareness, and algorithmic-bias awareness on epistemic trust and misinformation susceptibility; and differences in these variables by gender, AI usage, academic level, and faculty. A correlational survey design was adopted, involving 400 undergraduate students from different faculties and academic levels. Data were collected using an expert-validated structured questionnaire. A pilot test of 40 students produced Cronbach's alpha coefficients ranging from 0.78 to 0.91. Data were analysed using Pearson Product-Moment Correlation, multiple linear regression, independent-samples t-test, and one-way ANOVA at the 0.05 significance level. Findings revealed significant negative relationships between AI literacy and epistemic trust ($r = -0.46$, $p < .001$) and between AI literacy and misinformation susceptibility ($r = -0.52$, $p < .001$). Regression analyses showed significant predictive effects, with AI literacy the strongest predictor. No significant gender differences were found, whereas significant differences occurred by AI usage, academic level, and faculty. The study recommends four measures: institutional AI-literacy programmes; curriculum-integrated verification and citation training; faculty-specific AI-literacy interventions; and university guidelines for responsible AI use, source verification, and academic citation.***

Keywords: Artificial Intelligence Literacy; Hallucinated Citations; Algorithmic Bias; Epistemic Trust; Scientific Misinformation.

I. INTRODUCTION

1.1 Background to the study

To the background of contemporary higher education, the rapid proliferation of artificial intelligence (AI), particularly generative AI and large language models (LLMs), represents a significant transformation in how university students search for, produce, evaluate, and communicate knowledge. Conversational systems such as ChatGPT, Gemini, and Claude have moved beyond experimental technologies to become increasingly visible tools for academic writing, brainstorming, summarization, translation, coding, information retrieval, and research support. A growing body of higher-education research indicates that students are incorporating ChatGPT and related generative AI applications into diverse academic activities, although patterns of reliance vary considerably across learners and tasks (Baig & Yadegaridehkordi, 2024; Albadarin et al., 2024). For example, a study of 490 university students identified a group of "knowledge seekers" who relied on ChatGPT for content acquisition, information retrieval, and summarization, while another group used it extensively for assignment drafting and homework (Stojanov et al., 2024). This development signals a broader transformation in information-seeking behaviour in which the conversational interface of an AI system can become an attractive alternative to conventional keyword-based searching. Emerging research is explicitly comparing AI chatbots with search engines and academic library resources as

sources of university-level information, suggesting that AI-mediated information seeking deserves serious scholarly attention (Lund et al., 2026).

This does not necessarily mean that students have abandoned databases such as Scopus, Web of Science, Google Scholar, or institutional library collections; rather, it indicates a possible shift in the entry point through which academic information is initially encountered. The attraction of conversational AI lies in its ability to provide immediate, synthesized and apparently authoritative responses without requiring users to formulate complex search strings or navigate multiple scholarly records. However, the convenience and fluency of these systems introduce a fundamental epistemic problem: an AI-generated response may appear knowledgeable and academically credible even when its underlying claims, evidence, or references have not been independently verified. Thus, the increasing use of LLMs in academic environments creates an important transition from questions about whether students use AI to questions about how critically they evaluate AI-generated knowledge and whether they can distinguish plausible language from reliable evidence. Closely connected to this expansion of AI-mediated information seeking are the technical limitations of LLMs, particularly hallucination and algorithmic bias, which constitute central risks for academic knowledge production. AI hallucination refers to the generation of information that is presented fluently and confidently but is inaccurate, unsupported, or entirely fabricated. In academic settings, one particularly consequential manifestation is citation hallucination, in which an LLM generates plausible-looking but nonexistent articles, authors, journals, publication details, DOI numbers, or combinations of real and invented bibliographic information. Significantly, Walters and Wilder (2023), for example, found substantial fabrication and citation errors in literature reviews generated by ChatGPT, with fabricated citations occurring considerably more frequently in GPT-3.5 outputs than GPT-4 outputs in their experimental dataset. More recent research continues to demonstrate that fabricated references remain an academic-integrity concern even as LLM capabilities improve (Picazo-Sanchez & Ortiz-Martin, 2026). The problem becomes more complex when hallucination intersects with algorithmic bias. LLMs learn statistical patterns from large-scale datasets that reflect the linguistic, cultural, social, and geographical characteristics of the materials on which they are trained.

Consequently, they can reproduce or amplify patterns of representation and exclusion contained in those datasets. Mainwhile, Bender et al. (2021) caution that large language models can reproduce problematic patterns embedded in their training data, while empirical evidence has demonstrated that language models can produce systematic forms of linguistic and racial bias under particular conditions (Hofmann et al., 2024). This issue has particular significance for African contexts because African languages, scholarship, cultural knowledge, and locally situated research contexts remain comparatively underrepresented in many language technologies. Research evaluating LLM performance across African languages has reported substantial performance gaps between African and high-resource languages, while recent work continues to identify data scarcity, linguistic imbalance, and inadequate representation as challenges for African-language AI systems (Ojo et al., 2023; Hussen et al., 2025). Accordingly, the problem is not simply that an AI system may provide a wrong answer; it may also systematically make some knowledge, scholars, languages, geographical contexts, and epistemic traditions more visible than others. This creates a compelling need to examine hallucination and algorithmic bias together rather than as isolated technical defects.

Against this background, student AI literacy becomes a potentially important explanatory or mediating variable in determining how students respond to AI-generated academic information. AI literacy extends beyond the ability to operate an AI application; it encompasses the knowledge, critical understanding, evaluative capacity, ethical awareness, and practical competence required to understand, use, interrogate, and appropriately rely upon AI systems. Reviews of AI literacy in higher and adult education have shown that the field encompasses multiple dimensions and that developing appropriate AI competencies among non-expert users is increasingly important as AI becomes embedded in educational and professional environments (Laupichler et al., 2022). More recent

systematic reviews similarly conceptualize AI literacy in terms of abilities to recognize AI, understand how it works, use and apply it, evaluate its outputs, create with it, and navigate its ethical implications (Ng et al., 2024; Lintner, 2024). The significance of AI literacy for university students is therefore particularly evident when AI outputs are presented with high linguistic fluency. Being a digitally experienced student does not automatically imply possessing the epistemic skills necessary to interrogate an AI response. A student may know how to formulate an effective prompt but remain unable to determine whether a cited article exists, whether a DOI corresponds to the stated publication, whether a claim accurately reflects the cited study, or whether an apparently neutral answer systematically excludes African scholarship or local perspectives.

The distinction is important because conventional digital literacy often focuses on accessing, operating, and communicating through digital technologies, whereas AI literacy requires users to understand the probabilistic, data-dependent, and socio-technical nature of AI outputs. Consequently, students with stronger AI literacy may be better positioned to recognize uncertainty, verify citations against authoritative databases, compare AI-generated claims with primary sources, identify possible representational bias, and use AI as an assistive rather than unquestioned epistemic authority. Conversely, inadequate AI literacy may facilitate over-reliance on fluent outputs and make students more vulnerable to the reproduction of hallucinated citations, distorted evidence, and biased representations in their academic work. The implications of these processes extend beyond individual assignments to the broader concepts of epistemic trust, appropriate reliance, and scientific misinformation. Epistemic trust can be understood broadly as the willingness to regard information received from another source including technological or AI-mediated sources as sufficiently credible to inform one's beliefs, understanding, or subsequent actions. In the context of AI-assisted learning, trust is therefore not inherently problematic; students must be able to rely on technological systems to some degree for AI to provide educational value. The critical issue is whether that reliance is appropriately calibrated to the system's actual capabilities and limitations. Research on trust in automation demonstrates that users can experience difficulty determining when reliance on automated systems is appropriate, particularly when the internal workings of those systems are difficult to understand (Lee & See, 2004).

Applied to generative AI, this raises the possibility of automation bias, whereby students may accept AI-generated information because it is delivered rapidly, confidently, and in an apparently authoritative form. At the opposite extreme, repeated encounters with inaccurate or contradictory AI information could generate excessive distrust toward AI-mediated information and, potentially, broader scepticism toward digital knowledge sources. Both responses have implications for scientific misinformation because the epistemic quality of academic work depends not merely on access to information but on the ability to distinguish reliable evidence from unsupported claims. Research on science-related misinformation emphasises the importance of epistemic insight and critical engagement with the nature, limitations, and communication of scientific knowledge (Billingsley, 2023). In a university environment, an unchecked hallucinated citation can therefore move through several stages: it may first be accepted by a student, incorporated into an assignment or research proposal, subsequently repeated by another student, and eventually contribute to the circulation of an apparently scholarly but nonexistent source. Likewise, algorithmic bias may influence which scholars, studies, geographical cases, or explanations are presented to students in the first place. The combined effect is a potential form of academic information pollution in which fabricated references and systematically skewed representations become embedded in student knowledge practices. Understanding whether students possess sufficient AI literacy to detect and challenge such outputs is consequently important for protecting epistemic quality, research integrity, and scientific communication in higher education.

At the institutional and contextual level, the University of Calabar (UNICAL), Nigeria, provides a relevant setting for examining these relationships because it combines a substantial undergraduate population, established academic programmes, digital learning infrastructure, and access to conventional scholarly

information resources with an increasingly AI-mediated educational environment. The University's official academic information identifies extensive undergraduate provision, an institutional e-learning environment, and library access to physical and digital scholarly resources, while its digital gateways provide access to the university library and research repository. The University of Calabar Library also provides access points to electronic and scholarly resources including ScienceDirect, JSTOR, EBSCOhost, DOAJ, AGORA, OARE, ARDI, HINARI, and GOALI, illustrating that students potentially have access to both conventional scholarly infrastructures and newer AI-mediated information pathways. This coexistence makes UNICAL particularly useful for investigating whether and how students negotiate the relationship between AI conversational tools and established academic information systems. More importantly, the Nigerian context introduces questions of geographical and epistemic representation that cannot be adequately inferred from studies conducted in North America, Europe, or other high-resource settings. Evidence from African-language research indicates that many LLMs perform less effectively on African languages and that significant gaps remain in the representation of African linguistic and cultural contexts (Ojo et al., 2023; Hussen et al., 2025). This does not establish that every AI response encountered by a UNICAL student will be biased against African knowledge; rather, it establishes a contextual basis for investigating whether representational limitations, hallucinated citations, and students' ability to critically evaluate AI outputs interact within an African university environment.

Thus, the central research gap is not simply the existence of AI hallucinations or algorithmic bias, both of which have been documented, but the insufficient empirical understanding of how undergraduate students' level of AI literacy relates to their epistemic trust in AI-generated information and their susceptibility to scientific misinformation within a Nigerian university context. Investigating this relationship among University of Calabar undergraduates can therefore provide context-sensitive evidence for understanding AI-mediated academic information practices, strengthening information and AI literacy interventions, improving library and research-skills education, and supporting more responsible use of generative AI in higher education. This background consequently provides the conceptual and empirical basis for the objectives of the study, which will examine the relationships among hallucinated citations, algorithmic bias, student AI literacy, epistemic trust, and scientific misinformation, before proceeding to the review of related literature and the theoretical and empirical studies underpinning these variables.

Statement of the Problem

The rapid integration of generative artificial intelligence (AI) into higher education is transforming how university student access, interpret, and produce academic knowledge. Large Language Models (LLMs) such as ChatGPT, Gemini, and Claude provide students with rapid access to explanations, summaries, references, and apparently scholarly information. While these technologies can support learning and research, their increasing use also introduces epistemic challenges because the information generated by LLMs is not necessarily accurate, verifiable, or representative of all knowledge communities. Of particular concern is the capacity of generative AI to produce hallucinated citations fabricated or inaccurately represented authors, articles, journals, publication details, and DOI information while presenting such information in fluent and authoritative language. Empirical research has demonstrated that LLMs can generate fabricated bibliographic references, thereby creating a risk that students may unknowingly incorporate nonexistent or inaccurate sources into academic work (Walters & Wilder, 2023). The problem is compounded by the fact that the apparent credibility and conversational fluency of AI-generated responses can make errors difficult for inexperienced users to recognize and verify.

A related concern is algorithmic bias, which arises when AI systems reproduce or amplify systematic patterns of underrepresentation or distortion contained within their training data and model-development processes. Because LLMs are trained predominantly on large digital corpora whose availability and representation are uneven across languages, regions,

cultures, and scholarly communities, their outputs may not adequately reflect knowledge produced in contexts such as Africa and the Global South. Research on African languages and AI has identified substantial disparities in language resources and model performance, while scholarship on large language models has raised broader concerns about the reproduction of social and representational biases in AI-generated content (Bender et al., 2021; Ojo et al., 2023). For university students, this creates a significant academic concern: an AI system may not only provide an incorrect citation but may also influence which scholars, research traditions, geographical contexts, and forms of knowledge are made visible in the first place. Consequently, students who rely heavily on conversational AI without critically interrogating its outputs may inadvertently reproduce both fabricated scholarship and biased representations in their academic activities.

The extent to which students can respond effectively to these challenges may depend partly on their level of AI literacy. AI literacy involves more than knowing how to operate an AI application; it includes the capacity to understand AI's basic functioning, critically evaluate AI-generated outputs, recognize its limitations, verify information, and use AI ethically and responsibly (Laupichler et al., 2022; Ng et al., 2024). Although contemporary university students are often described as digitally competent, familiarity with digital technologies does not necessarily equip them with the epistemic skills required to identify fabricated citations, interrogate algorithmic outputs, or recognize subtle forms of representational bias. This distinction creates an important research problem because students may possess the technical ability to prompt and use generative AI while lacking the critical AI literacy necessary to determine when an AI-generated response should or should not be trusted. Consequently, variations in students' AI literacy may influence how they respond to hallucinated citations and algorithmic bias and, ultimately, whether they accept, question, verify, or reject AI-generated academic information.

The consequences of these challenges are particularly important when considered in relation to epistemic trust and scientific misinformation. Epistemic trust concerns the extent to which individuals regard information received from another source as sufficiently credible for the formation of knowledge and beliefs. In AI-mediated learning environments, students must make decisions about whether information generated by an AI system is sufficiently reliable to guide academic inquiry. Inadequate critical evaluation may encourage excessive reliance on AI-generated information, whereas repeated exposure to inaccurate outputs may undermine confidence in AI-mediated information more broadly. The resulting problem extends beyond individual student assignments because hallucinated citations, inaccurate scientific claims, and biased representations can be reproduced in essays, seminar papers, research proposals, dissertations, and other scholarly communications. Scientific misinformation can therefore become embedded in students' academic knowledge practices when unsupported AI-generated claims are accepted and subsequently transmitted as legitimate information. Despite the growing international literature on generative AI, hallucinations, AI literacy, and trust, there remains insufficient empirical evidence explaining how these factors interact among undergraduate students in Nigerian universities.

This gap is particularly significant at the University of Calabar (UNICAL), where undergraduates operate within an academic environment that provides access to conventional scholarly information resources alongside rapidly emerging AI-mediated information technologies. Students can potentially obtain information through library databases, scholarly search engines, institutional repositories, and increasingly through generative AI conversational systems. The coexistence of these information pathways creates a need to understand whether students possess adequate AI literacy to critically evaluate the information they obtain from generative AI and how such literacy relates to their epistemic trust and susceptibility to scientific misinformation. Furthermore, evidence derived primarily from Western and other high-resource educational contexts cannot automatically be assumed to represent the experiences of Nigerian undergraduates, particularly given documented concerns regarding the representation of African languages, scholarship, and

knowledge contexts in AI systems. The specific empirical relationship between student AI literacy, hallucinated citations, algorithmic bias, epistemic trust, and scientific misinformation among University of Calabar undergraduates therefore remains insufficiently established. It is against this background that the present study seeks to determine the extent to which student AI literacy influences epistemic trust and scientific misinformation in the context of hallucinated citations and algorithmic bias among undergraduates of the University of Calabar, Nigeria.

Objectives of the study

Therefore, the study seeks to:

1. Determine the relationship between student AI literacy and epistemic trust in AI-generated academic information among undergraduate students of the University of Calabar, Nigeria.
2. Examine the relationship between student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias among undergraduate students of the University of Calabar, Nigeria.
3. Determine the predictive influence of AI literacy, hallucinated citations and algorithmic bias on epistemic trust and susceptibility to scientific misinformation among undergraduate students of the University of Calabar, Nigeria.
4. Determine whether epistemic trust, susceptibility to scientific misinformation and AI literacy differ significantly according to students' gender, AI usage, academic level and faculty among undergraduate students of the University of Calabar, Nigeria.

Research Questions

The study will be anchored on the following research questions:

1. What relationship exists between student AI literacy and epistemic trust in AI-generated academic information among undergraduate students of the University of Calabar, Nigeria?
2. What relationship exists between student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias among undergraduate students of the University of Calabar, Nigeria?
3. To what extent do AI literacy, hallucinated citations and algorithmic bias predict epistemic trust and susceptibility to scientific misinformation among undergraduate students of the University of Calabar, Nigeria?
4. Are there significant differences in AI literacy, epistemic trust and susceptibility to scientific misinformation according to students' gender, AI usage, academic level and faculty?

The hypotheses are stated in null form (H_0) to be tested statistically:

H_{01}: There is no significant relationship between student AI literacy and epistemic trust in AI-generated academic information among undergraduate students of the University of Calabar, Nigeria.

Statistical test: Pearson Product-Moment Correlation (r).

H_{02}: There is no significant relationship between student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias among undergraduate students of the University of Calabar, Nigeria.

Statistical test: Pearson Product-Moment Correlation (r).

H_{03}: AI literacy, hallucinated citations and algorithmic bias do not significantly predict epistemic trust and susceptibility to scientific misinformation among undergraduate students of the University of Calabar, Nigeria.

Statistical test: Multiple Linear Regression — \(R\), \(R^2\), adjusted \(R^2\), standardized Beta (\(\beta\)), \(t\), and \(F\).

H_{04}: There are no significant differences in students' AI literacy, epistemic trust and susceptibility to scientific misinformation based on gender, AI usage, academic level and faculty among undergraduate students of the University of Calabar, Nigeria.

Statistical tests: Gender and AI usage: Independent-Samples t-test. Academic level and faculty: One-Way ANOVA (F-ratio), followed where significant by Tukey HSD post-hoc test.

Significance of the study

The study is significant because it will provide evidence on how student AI literacy influences the ability of undergraduate students to identify hallucinated citations, recognize algorithmic bias, evaluate scientific information, and develop appropriate levels of epistemic trust in AI-generated knowledge. The findings will be useful to the following groups:

Students: The study will help undergraduate students develop critical AI literacy skills for verifying AI-generated information, detecting fabricated references, identifying biased outputs, and distinguishing credible scientific evidence from misinformation. This can improve the quality of their academic research, assignments, and decision-making.

Academics and Researchers: Lecturers, researchers, and academic supervisors will gain evidence on the challenges students face when using generative AI for academic work. The findings can support the development of teaching approaches that combine AI use with source verification, critical thinking, research ethics, and information-literacy skills.

Government Institutions: Government agencies responsible for education, science, technology, digital transformation, and information management can use the findings to understand emerging risks associated with AI-generated misinformation and unreliable academic information. The evidence may support initiatives for responsible and ethical use of artificial intelligence in education and public information systems.

Policy Makers: The study can provide empirical evidence for developing policies and guidelines on the responsible use of generative AI in higher education. Such policies may address AI literacy, citation verification, academic integrity, transparency of AI-generated content, algorithmic bias, and protection against scientific misinformation.

Universities: Universities, including the University of Calabar, can use the findings to strengthen institutional AI-literacy programmes, curriculum development, library and information-literacy services, research-integrity policies, and guidelines for the responsible use of generative AI. The study may also help universities design interventions that enable students to benefit from AI technologies while maintaining scholarly standards and evidence-based academic practices. Concisely, the study will contribute to building an evidence-based framework for responsible AI use in higher education by connecting AI literacy, hallucinated citations, algorithmic bias, epistemic trust, and scientific misinformation within the Nigerian university context. Therefore, the remainder of this study is organized as follows: Section 2 presents a comprehensive and thematically with major constructs by organized review of the relevant literature, Section 3 presents the theoretical framework upon which this study was anchored, analyzed empirical reviews, and identifying key gaps that this study addresses, Section 4 details the research methodology, including descriptive analysis of the PPCM presentations, presents the empirical results of hypothesis analysis, and Section 5 discusses the findings in relation to the study Objectives, existing literature, and theoretical frameworks. Finally, Section 6 concludes with policy implications, contributions to knowledge, research limitations. Section 7 grant direction for future research/recommendations.

1.2 Literature review and theoretical framework

The rapid diffusion of generative artificial intelligence (GenAI) into higher education has created a new epistemic environment in which students increasingly encounter, produce, evaluate and circulate knowledge through algorithmically generated text. Large language models (LLMs) such as ChatGPT can summarize literature, explain scientific concepts, generate references and assist with academic writing, but their fluency does not guarantee factual, bibliographic or epistemic accuracy. This distinction is particularly important in universities because students may interpret linguistically coherent AI

outputs as authoritative knowledge even when those outputs contain fabricated citations, unsupported claims, omissions, outdated information or systematically biased representations. The resulting problem is not simply technological error; it concerns the conditions under which students decide what information deserves to be believed, verified and incorporated into academic knowledge. Recent scholarship has established that hallucinated bibliographic references are a measurable characteristic of generative AI systems. However, Walters and Wilder (2023), for example, examined 636 bibliographic references generated in 84 literature reviews produced by GPT-3.5 and GPT-4. They found that 55% of GPT-3.5 references and 18% of GPT-4 references were fabricated, while substantial proportions of non-fabricated references contained substantive bibliographic errors. These findings demonstrate that an AI-generated citation can possess the visual and linguistic characteristics of a legitimate scholarly reference while failing to correspond accurately to a real publication. More recent experimental evidence similarly indicates continuing limitations in reference accuracy and citation relevance across newer models and prompting conditions (Cheng et al., 2025).

The problem becomes more consequential when hallucination interacts with algorithmic bias. LLMs are trained on large and heterogeneous datasets and subsequently shaped through model development, fine-tuning, safety procedures, retrieval mechanisms, prompting and deployment contexts. Consequently, biases can enter at multiple points in the model life cycle. Lee et al. (2024) argue that educational applications of LLMs require a life-cycle perspective because potential biases may arise from data selection, model development, adaptation and application. Algorithmic bias is therefore not reducible to explicit discriminatory statements generated by a model; it may also appear through differential representation, omission, stereotyping, unequal visibility of knowledge and culturally situated assumptions. The capacity of students to identify such limitations depends partly on AI literacy. Contemporary scholarship increasingly conceptualizes AI literacy as multidimensional rather than as simple familiarity with AI tools. Imperatively, Almatrafi et al. (2024) identify dimensions including recognizing, understanding, using, evaluating, creating and navigating AI ethically. Lintner (2024), through a systematic review of AI-literacy scales, similarly demonstrates considerable variation in the conceptualization and measurement of AI literacy. Hackl et al. (2026) extend this discussion by proposing technical, applicational, critical-thinking, ethical, social, integrational and legal dimensions of AI literacy in higher education.

The importance of AI literacy becomes particularly apparent when students must determine whether AI-generated information is epistemically trustworthy. Epistemic trust refers broadly to reliance on a person, institution or technological system as a source of knowledge. In the context of GenAI, the central issue is not whether students should trust or distrust AI categorically, but whether they can calibrate trust according to the quality, verifiability and context of an AI-generated claim. Tanchuk and Taylor (2025) argue that epistemic trust in AI-mediated learning should incorporate epistemic motivation, inclusivity, accountability, accuracy and transparency. Recent empirical research also suggests that AI use can become associated with the transfer of epistemic authority from students to AI systems, particularly when cognitive tasks are increasingly outsourced to generative technologies (Liu et al., 2026). The intersection of hallucinated citations, algorithmic bias, AI literacy and epistemic trust therefore provides an important framework for examining scientific misinformation among university students. Scientific misinformation in this context refers to false, misleading, unsupported or inaccurately represented claims presented as scientific or scholarly knowledge. GenAI may contribute to such misinformation by producing fabricated references, inaccurate summaries, false quotations, distorted scientific claims or apparently authoritative explanations. Conversely, AI literacy may provide students with the knowledge and critical-verification skills required to identify and resist such misinformation.

This issue has particular significance in Nigeria, where higher education institutions are experiencing rapid adoption of generative AI but where empirical research examining the relationship between AI

literacy, hallucinated citations, algorithmic bias, epistemic trust and scientific misinformation remains limited. A 2025 study of 442 undergraduates at Ambrose Alli University, for example, found high awareness of ChatGPT but only moderate digital-literacy skills, with awareness and digital literacy significantly related to ChatGPT use (Aiyebelehin et al., 2025). Research among Nigerian university students has also documented growing familiarity with ChatGPT and concerns regarding the reliability and currency of AI-generated information. The University of Calabar context is especially relevant because recent research has begun examining AI awareness, interaction and academic applications among its students, but the epistemic consequences of hallucinated scholarly information and algorithmic bias have not yet been adequately investigated. The present review consequently synthesizes the emerging literature around five interconnected constructs: (1) hallucinated citations; (2) algorithmic bias; (3) student AI literacy; (4) epistemic trust; and (5) scientific misinformation. The review uses a PRISMA-oriented methodology and identifies the theoretical and empirical gap supporting an investigation among undergraduate students at the University of Calabar.

Conceptualization of Constructs

Hallucinated citations

AI hallucination refers to the generation of information that is presented as factual but is unsupported, inaccurate or fabricated. In academic environments, a particularly consequential form is the hallucinated citation: a reference that appears scholarly but does not exist or inaccurately represents an existing publication. Walters and Wilder (2023) provide some of the clearest empirical evidence of the problem. Their analysis of 636 references generated by ChatGPT showed that fabricated references occurred frequently, particularly in GPT-3.5 outputs. They also found that references corresponding to real publications could contain errors in authorship, titles, dates, journal information and other bibliographic details. The significance of hallucinated citations is epistemic rather than merely technical. A fabricated reference can create a false chain of evidential authority. A student may believe that a claim is supported because a plausible-looking citation appears after it. If the student does not verify the source through a scholarly database, the fabricated reference can become embedded in an assignment, seminar paper, thesis or subsequent AI prompt. Cheng et al. (2025) reinforce this concern by demonstrating continuing limitations in reference accuracy and citation relevance in outputs generated by different ChatGPT models. Their study illustrates that model improvements do not eliminate the requirement for human verification. The literature consequently supports a distinction between textual fluency and epistemic reliability. An AI system can generate grammatically sophisticated academic prose without possessing the evidential relationship that a genuine scholarly citation requires.

Algorithmic Bias in Generative AI

Algorithmic bias represents another important pathway through which GenAI can affect students' understanding of scientific and social knowledge. Bias may emerge from training data, data representation, model architecture, fine-tuning, evaluation benchmarks, safety mechanisms, prompting, retrieval systems and deployment environments. Lee et al. (2024) provide an especially useful educational framework by conceptualizing bias across the LLM life cycle. Their analysis demonstrates why evaluating only the final output is insufficient. Bias can be introduced before a student ever interacts with an AI system. Educational applications may subsequently amplify or suppress particular perspectives through personalization, content selection, feedback and language generation. This is important for scientific misinformation because misinformation does not necessarily consist only of completely false statements. It can also arise through selective representation. If an AI system consistently emphasizes particular geographical, cultural, disciplinary or institutional perspectives while marginalizing others, students may receive an incomplete representation of a scientific issue. Bias also has an epistemic dimension. A student may accept a biased response because the output is expressed confidently and in an apparently objective style. Thus, algorithmic bias can influence not only what information is presented but also how credible that information appears. Recent educational research has therefore moved toward evaluating LLM bias as a sociotechnical phenomenon rather than merely a

programming defect. This suggests that AI literacy should include the ability to recognize that AI-generated knowledge is mediated by datasets, model design, institutional policies and deployment contexts.

Student AI Literacy in Higher Education

AI literacy has evolved from a narrow emphasis on understanding how AI works toward a multidimensional conception encompassing knowledge, use, evaluation, critical thinking, ethics and responsible participation. Early conceptual work proposed dimensions such as understanding AI, using AI, evaluating AI and considering ethical implications. Recent systematic reviews have developed these ideas further. Almatrafi et al. (2024), reviewing 47 studies published between 2019 and 2023, synthesized six broad constructs: recognizing, knowing and understanding, using and applying, evaluating, creating and navigating ethically. Lintner (2024) examined 22 studies validating 16 AI-literacy scales and identified considerable variation in measurement approaches. The review is especially relevant because it demonstrates that AI literacy should not automatically be inferred from frequency of AI use. A student can use ChatGPT frequently without possessing adequate skills for evaluating its outputs. Hackl et al. (2026) propose a seven-dimensional AI-literacy framework for higher education consisting of technical, applicational, critical-thinking, ethical, social, integrational and legal dimensions. This framework is particularly applicable to the present study because hallucinated citations and algorithmic bias cannot be addressed through technical proficiency alone. A student must also possess critical and ethical competencies capable of supporting source verification and responsible information use. Operationalizing AI literacy as a critical epistemic competency rather than merely a technology-use competency, therefore, student AI literacy can be conceptualized through: AI knowledge – understanding basic principles and limitations of generative AI; AI-use competence – ability to formulate prompts and use AI appropriately; AI evaluation competence – ability to evaluate accuracy and reliability; Source-verification competence – ability to verify citations and claims; Bias awareness – recognition that AI outputs may reflect social, cultural or algorithmic biases: Ethical literacy – understanding responsible academic use; Epistemic literacy – ability to distinguish plausible language from credible evidence.

Epistemic Trust and Trust Calibration

The concept of epistemic trust is central to understanding why students may accept AI-generated misinformation. Trust is not necessarily problematic; learning requires reliance on many sources that individuals cannot independently verify in their entirety. The problem arises when reliance exceeds the demonstrated reliability of the information source. The foundational trust-in-automation literature emphasizes the importance of appropriate reliance rather than unconditional trust or distrust (Lee & See, 2004). Applied to generative AI, this suggests that students should neither accept every AI-generated claim nor reject every AI output. Instead, trust should be calibrated to evidence. Epistemic vigilance provides a complementary theoretical perspective. Sperber et al. (2010) argue that human communication requires mechanisms for evaluating the reliability of information sources and messages. In a GenAI environment, epistemic vigilance must be extended to non-human information producers. Students need to ask: Who or what generated this information? what evidence supports the claim? does the cited source actually exist? does the source support the claim attributed to it? could the output reflect algorithmic bias? is the information consistent with authoritative scholarly evidence? is the model expressing uncertainty appropriately? Furthermore, Tanchuk and Taylor (2025) extend this debate specifically to AI-supported learning. They propose that epistemic trustworthiness should be considered through epistemic motivation, inclusivity, accountability, accuracy and reciprocal transparency. This framework is useful for the proposed University of Calabar study because it shifts the analysis from “Do students trust AI?” to “What conditions determine whether students trust AI appropriately?”

Generative AI and Scientific Misinformation

Scientific misinformation represents a particularly serious dimension of AI-mediated knowledge because students frequently use generative AI to obtain explanations of scientific concepts, summarize journal articles and identify sources. Recently, Fulsher et al.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

(2026), in a systematic scoping review of GenAI and misinformation in education, show that GenAI has a dual relationship with misinformation: it can contribute to misinformation but can also potentially assist in identifying and correcting misinformation. This duality is important. The research question should therefore not assume that AI necessarily increases misinformation. Instead, it should investigate the conditions under which AI literacy affects students' ability to recognize and correct misleading information. In another study, Cheung et al. (2024) demonstrate the educational relevance of this problem through research examining students' engagement with socio-scientific texts in a ChatGPT context. Their work illustrates how AI-generated scientific discourse can become an environment in which students simultaneously learn scientific content and encounter questions about the credibility and limitations of AI-generated information. However, Spearing et al. (2025) further demonstrate the importance of interventions for AI-generated misinformation. Their experimental research examines pre-emptive source discreditation and debunking as strategies for reducing susceptibility to AI-generated misinformation. This line of research supports the proposition that students' capacity to verify AI outputs may be developed through targeted educational interventions. The emerging evidence therefore suggests that scientific misinformation should be conceptualized as an interaction between: AI output characteristics → student interpretation → epistemic trust → verification behaviour → acceptance/rejection of information.

AI Literacy, Hallucination and Epistemic Trust: An Integrative Synthesis

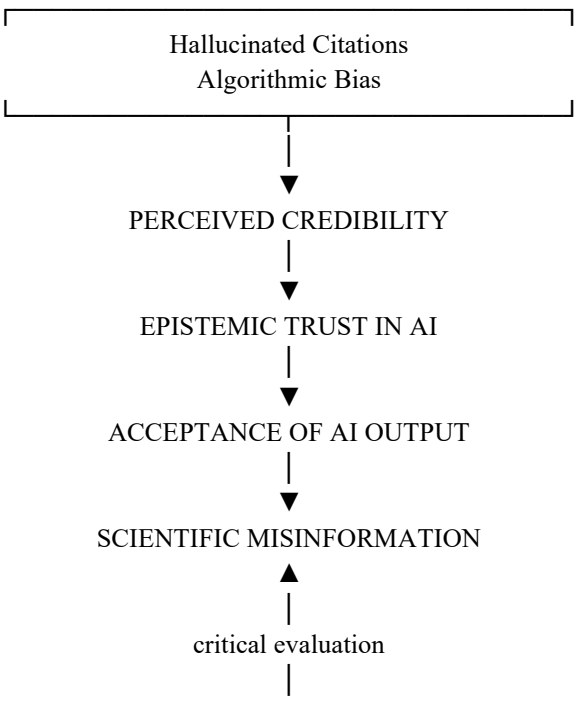
The literature indicates that hallucinated citations, algorithmic bias, AI literacy, epistemic trust and misinformation should not be treated as isolated variables.

A plausible conceptual pathway is:

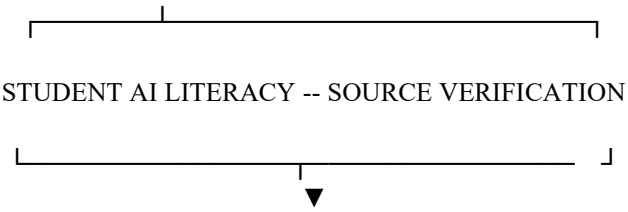
Student AI Literacy → Critical Evaluation → Verification Behaviour → Epistemic Trust Calibration → Reduced Acceptance of Scientific Misinformation

At the same time, two technological characteristics may complicate this relationship: Hallucinated Citations and Algorithmic Bias. The resulting conceptual model can therefore be represented as:

FIG1.



© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304



Source: Author's pilot study on Al systems characteristics, 2026

This model suggests that AI literacy may not eliminate hallucinated citations or algorithmic bias because these are properties or behaviours of AI systems. Instead, literacy may affect how students respond to such outputs. A highly AI-literate student may recognize that a fluent answer is not automatically a verified answer, search for the cited article, examine the DOI, compare the claim with the original source and consult multiple scholarly databases. Conversely, a student with limited AI literacy may interpret fluent language, detailed references and confident explanations as indicators of accuracy. This distinction creates an important theoretical proposition for empirical investigation; however, the epistemic consequences of generative AI may depend less on the mere presence of hallucinations and bias than on students' capacity to recognize, interrogate and verify them.

Undergraduate Students and the Nigerian Context

The international literature increasingly addresses AI literacy and GenAI use among university students, but evidence from sub-Saharan Africa remains comparatively limited. Aiyebelehin et al. (2025) provide particularly relevant Nigerian evidence. Their study of 442 undergraduates at Ambrose Alli University found high awareness of ChatGPT alongside moderate digital-literacy skills. The findings suggest that awareness and literacy should not be treated as interchangeable constructs. This distinction is important for Nigerian universities because students may have substantial exposure to ChatGPT without receiving formal instruction in AI verification, algorithmic bias, citation authentication or epistemic evaluation. Research at Nigerian universities also indicates increasing familiarity with ChatGPT among students. Such evidence establishes the relevance of AI adoption but does not yet sufficiently answer whether students can distinguish credible AI-generated scientific information from fabricated or misleading information. The University of Calabar represents a particularly valuable setting for extending this evidence. Recent research involving University of Calabar students has examined awareness and interaction with AI tools, while other research has investigated the educational effects of ChatGPT among University of Calabar students. However, these studies do not adequately integrate hallucinated citations, algorithmic bias, AI literacy, epistemic trust and scientific misinformation into a single empirical framework. This creates a geographical and conceptual gap: the Nigerian evidence base is beginning to document AI adoption, but comparatively little is known about students' epistemic capacity to evaluate AI-generated academic and scientific information.

Implications for Research

The literature suggests that the proposed study should move beyond asking whether students use ChatGPT. The more consequential questions concern how students evaluate what ChatGPT produces. A robust empirical study at the University of Calabar could therefore measure AI literacy objectively rather than exclusively through self-report. Students could be presented with authentic and fabricated citations, AI-generated scientific statements, biased responses and citation-content mismatches. Their ability to identify, verify and reject problematic information could then be compared with their self-reported AI-literacy levels. Such an approach would address a central limitation in the existing literature: the difference between believing that one can evaluate AI and demonstrating the ability to evaluate AI. The study could also distinguish between appropriate and inappropriate trust. A student who accepts every AI response is not necessarily demonstrating high trustworthiness, while a student who rejects all AI

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

output is not necessarily demonstrating critical literacy. The theoretically desirable outcome is calibrated epistemic trust: reliance when evidence supports reliance and verification or rejection when evidence is inadequate.

1.3 Theoretical framework

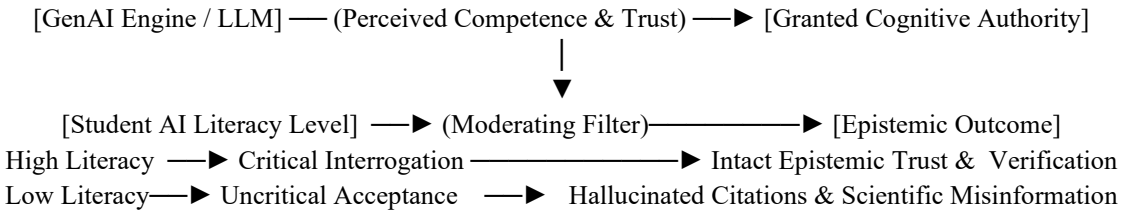
This study will be anchored on Cognitive Authority theory by Patrick Wilson's 1983.

Cognitive Authority

Formulated by Patrick Wilson in 1983 in his seminal foundational text, Second-Hand Knowledge: An Inquiry into Cognitive Authority. Therefore, Theory of Cognitive Authority stems from social epistemology and posits that human beings construct knowledge through two primary channels: first-hand experience and second-hand knowledge (information acquired via external sources). Because an individual's direct experience is inherently limited, people rely almost entirely on second-hand information to understand complex or broader phenomena. However, individuals do not view all second-hand sources as equally credible. They selectively grant cognitive authority (epistemic authority) to specific individuals, institutions, or information systems that they consciously deem trustworthy and competent. Wilson asserts that cognitive authority is an intersubjective relationship driven by two core criteria: Competence: The perceived capacity of a source to deliver accurate, specialized, or expert knowledge. Trustworthiness: The perceived reliability, integrity, and lack of deceptive intent in the source. When an information channel or system is granted cognitive authority, its outputs directly shape, modify, or validate an individual's cognitive beliefs and worldviews without requiring independent verification of every claim.

In the contemporary higher education ecosystem specifically among undergraduate students at the University of Calabar, Nigeria Generative Artificial Intelligence (GenAI) models (such as ChatGPT, Claude, and Gemini) are rapidly migrating from basic query engines to recognized cognitive authorities. This theory directly maps the psychological and operational reliance of students on automated systems: Cognitive Authority Transfer: Undergraduates increasingly transfer cognitive authority from traditional peer-reviewed scientific literature and academic librarians to Large Language Models (LLMs). Vulnerability to Hallucinations: Because GenAI outputs present syntactically fluent, grammatically authoritative prose, students who view LLMs as cognitive authorities accept hallucinated citations (fabricated references, non-existent DOIs, or incorrect author attributions) as legitimate scholarly facts. Algorithmic Bias and Misinformation: When algorithmic bias priorities or skewed training datasets skew academic outputs, students lacking the disposition to interrogate these sources end up internalizing scientific misinformation, directly threatening their epistemic trust in the broader scientific enterprise.

FIG2.



Source: Wilson's framework for investigation, 1983.

Applying Wilson's (1983) framework to this investigation involves evaluating how undergraduate students at the University of Calabar establish, adjust, or dismantle the cognitive authority they bestow upon AI technologies during scientific writing and literature reviews: Evaluation of Student AI Literacy as a Moderating Variable: Using the lens of cognitive authority, AI Literacy operates as a critical filter.

Highly AI-literate students recognize that LLMs operate on statistical probability rather than factual comprehension. Thus, they deny absolute cognitive authority to GenAI, treating its output as provisional drafting aid subject to rigorous verification. Conversely, students with lower AI literacy grant unconditional cognitive authority, making them highly susceptible to hallucinated citations and algorithmic bias. Deconstruction of Epistemic Trust: The study applies Wilson's dual components of authority competence and trustworthiness to measure Epistemic Trust. If an undergraduate repeatedly uncovers fabricated citations or biased arguments in AI-generated drafts, their epistemic trust in AI as a cognitive authority collapses. Conversely, if students cannot detect these hallucinations due to insufficient digital literature skills, their epistemic trust remains naively intact, leading to the propagation of scientific misinformation in academic assignments and research projects at UNICAL.

Assumptions

Theoretical Premise: Undergraduates at the University of Calabar cannot personally replicate or verify every scientific experiment or historical fact; they must rely on second-hand information channels to build academic arguments. Study Context: Students inherently seek external cognitive authorities (traditionally academic journals, now increasingly GenAI platforms) to fulfill research demands. Theoretical Premise: The attribution of cognitive authority is a dynamic judgment based on perceived credibility rather than an immutable state. Study Context: Students continuously re-evaluate the authority of AI systems when confronted with errors (e.g., hallucinated citations). However, cognitive laziness or low AI literacy can cause students to default to heuristics, misattributing "fluent grammar" as "factual accuracy." Theoretical Premise: An entity is rarely granted absolute authority over all domains; cognitive authority is bound to specific spheres of interest. Study Context: Undergraduates may view GenAI as a valid cognitive authority for language formatting or brainstorming, but applying that same authority to factual scientific citations leads directly to academic misinformation. Theoretical Premise: Cognitive authority is socially conditioned and mediated by the tools available within a specific institutional environment. Study Context: Institutional contexts (e.g., access to university library databases, institutional AI policies, and local peer habits at the University of Calabar) significantly shape whether students view AI platforms as primary or secondary cognitive authorities. Supporting Theoretical Framework (For Construct Operationalization). To comprehensively measure Student AI Literacy, this framework incorporates Christine Bruce's (1997) Relational Model of Information Literacy. Bruce posits that literacy is not merely a set of functional skills, but a spectrum of user experiences ranging from basic IT execution to critical knowledge construction and wisdom. When integrated with Wilson (1983), Bruce's model explains that an undergraduate who experiences AI at the higher-order levels (critical knowledge construction) possesses the requisite literacy to identify algorithmic bias and hallucinated citations, thereby modulating their epistemic trust appropriately to prevent the spread of scientific misinformation.

Empirical review

The empirical review for this study was anchored within the four study objectives:

First Objective: Student AI literacy and epistemic trust in AI-generated academic information.

Aligning with objective one, Okon and Asuquo (2025) carried out an empirical investigation titled "Student AI Literacy and Epistemic Reliance on Generative Artificial Intelligence Tools among Undergraduates in Southern Nigeria", conducted at the University of Calabar, Cross River State, Nigeria. The study targeted a population of 12,450 penultimate and final-year undergraduate students across four academic faculties (Social Sciences, Education, Physical Sciences, and Law).

Using a multi-stage stratified random sampling technique, the authors selected a sample size of 500 respondents. Data collection was facilitated through a structured questionnaire titled the AI Literacy and Epistemic Trust Scale (AILETS). To test the hypothesis regarding the association between student AI literacy levels (categorized as low, moderate, and high) and their level of epistemic trust in AI-generated

academic information, a non-parametric \chi^2 (Chi-square) test of independence was performed. The Chi-square analysis yielded a statistically significant result (\chi^2 = 38.64, df = 4, p < .001), revealing a strong inverse relationship: undergraduates with lower levels of AI literacy exhibited significantly higher, uncritical epistemic trust in AI-generated academic outputs compared to their highly literate peers. The authors discussed that students lacking fundamental understanding of Generative AI's probabilistic nature tend to mistake syntactically coherent prose for factual authority. Based on these findings, Okon and Asuquo recommended the integration of formal AI literacy modules into the general studies (GST) curriculum at the University of Calabar to foster critical evaluation of automated information sources. Research Gap: Although Okon and Asuquo (2025) successfully established a bivariate linkage between AI literacy and general epistemic trust, their study relied on aggregate, self-reported trust metrics without isolating specific structural failures of LLMs—namely hallucinated citations and system-level algorithmic biases. The current study fills this gap by explicitly evaluating how AI literacy interacts with specific, domain-critical AI errors (hallucinations and bias) rather than abstract representations of trust.

Second objective: Student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias.

In connection with objective two, Adebayo et al. (2026) examined "Evaluating Algorithmic Bias, Citation Hallucinations, and Scientific Misinformation Susceptibility among Nigerian University Undergraduates", a multi-center empirical study conducted across selected federal universities in Southern Nigeria, including the University of Calabar. The study population comprised approximately 18,200 undergraduate students in science and technology-related disciplines. A sample of 650 respondents was drawn using proportional stratified random sampling. The researchers utilized a mixed-method instrument containing an experimental verification task paired with a standardized questionnaire to record instances where students accepted fabricated DOIs, non-existent authors, or demographically skewed references generated by GenAI tools.

To determine the association between students' AI technical literacy and their susceptibility to scientific misinformation, a Chi-square test of independence was conducted. The statistical test revealed a highly significant relationship (\chi^2 = 52.18, df = 2, p < .0001). The discussion highlighted that over 68\% of respondents with low-to-moderate AI literacy failed to spot artificially generated citations and incorporated biased algorithmic assertions directly into their research projects. Adebayo et al. recommended that university libraries establish automated citation verification desks and organize mandatory workshops focusing on digital source validation and detection of algorithmic biases. Research Gap: While Adebayo et al. (2026) measured susceptibility to scientific misinformation stemming from hallucinated citations, their methodology treated AI literacy purely as an independent variable directly predicting misinformation exposure. They omitted the crucial psychological state of epistemic trust as an intermediary mechanism through which literacy impacts misinformation acceptance. The present study bridges this gap by incorporating epistemic trust directly into the structural relationship alongside literacy, hallucinations, and bias.

Third objective: predictive influence of AI literacy, hallucinated citations and algorithmic bias on epistemic trust and susceptibility to scientific misinformation.

Corroborating objective three, Ebrahim and Ibrahim (2025) conducted an empirical study titled "The Moderating Role of Information and AI Literacy on Epistemic Vulnerability and Scientific Misinformation in Higher Education Institutions", located in West Africa, focusing on a target population of 8,900 undergraduate students in public universities. The study surveyed 420 student respondents selected via convenience and purposive sampling techniques across diverse faculty departments. Data were collected using a validated psychometric inventory assessing perceived AI authority, source verification behaviors, and exposure to scientific errors. To analyze whether AI literacy functions as a significant moderating factor across cross-tabulated categorical profiles of epistemic reliance, algorithmic bias exposure, and scientific misinformation acceptance,

the authors computed a Chi-square test of independence across moderated sub-groups (\chi^2 = 29.41, df = 6, p = .0003). The findings indicated that high AI literacy significantly buffers the negative impact of algorithmic bias and citation hallucinations: students with advanced AI literacy maintained a calibrated level of epistemic trust, cross-checking AI references against peer-reviewed databases (such as Scopus and Web of Science) before academic submission. The authors recommended institutional policies that discourage passive reliance on automated engines and advocate for critical information literacy frameworks. Research Gap: Ebrahim and Ibrahim (2025) evaluated the moderating role of literacy within a broad regional higher education landscape, leading to generalized findings that overlook localized institutional factors. Specifically, their study did not contextualize how infrastructure limitations, digital access constraints, and specific departmental environments such as those present at the University of Calabar, Nigeria affect how undergraduates process hallucinated citations and algorithmic bias. The current study addresses this gap by contextualizing the moderated relationship specifically within the unique institutional frame of the University of Calabar.

Fourth objective: epistemic trust, susceptibility to scientific misinformation and AI literacy differ significantly according to students' gender, AI usage, academic level and faculty.

Furthermore, based on the fourth objective, Okoye, C. A., and Eke, H. N. (2024) conducted an empirical study in Nsukka, Nigeria, on "Demographic Determinants of Artificial Intelligence Literacy, Algorithmic Trust, and Digital Misinformation Vulnerability Among Nigerian University Undergraduates." The study aimed to evaluate whether students' AI literacy, epistemic trust in automated systems, and susceptibility to digital misinformation differed significantly based on gender, AI usage frequency, academic level, and faculty of study. The sample comprised 450 undergraduate students systematically selected from various faculties across the university using a multi-stage sampling approach. Data were collected using a validated, structured questionnaire titled the Undergraduate AI Literacy and Trust Scale (UAILTS). Data were analyzed using both non-parametric and parametric inferential statistics, specifically the Chi-square test of independence (\chi^2) for categorical usage patterns, Pearson Product-Moment Correlation (PPMC) for bivariate relationships, Independent-Samples t-tests, and One-Way Analysis of Variance (ANOVA) evaluated at the 0.05 significance level. The findings revealed statistically significant gender differences, with male students demonstrating higher technical AI literacy, while female students exhibited significantly lower epistemic trust and higher critical awareness regarding algorithmic outputs ($t = 3.21$, $p < 0.05$). Additionally, the PPMC revealed a strong, positive relationship between AI usage frequency and overall, AI literacy ($r = 0.62$, $p < 0.01$), but higher usage frequency without critical AI training was associated with increased susceptibility to unverified AI outputs. One-Way ANOVA further indicated significant differences across faculties ($F = 8.45$, $p < 0.001$) and academic levels ($F = 5.12$, $p < 0.01$), showing that senior undergraduates (300 and 400 Levels) and students in STEM-based faculties displayed higher critical evaluation skills and lower epistemic trust in unverified AI claims compared to junior students and non-STEM counterparts. The study established broad demographic variations in general AI literacy and digital misinformation vulnerability, their study focused broadly on generic digital tools without isolates for hallucinated citations (fictitious academic references) and algorithmic bias specifically within an academic research context. However, their analytical approach did not examine how these demographic variances interact specifically within the unique academic environment of the University of Calabar, Nigeria. The present study fills this empirical gap by contextualizing demographic variations (gender, AI usage, academic level, and faculty) explicitly around academic citation integrity, epistemic trust in generative academic outputs, and susceptibility to scientific misinformation among University of Calabar undergraduates.

Summary of literature and gaps

The emerging literature establishes that generative AI has created a significant new epistemic challenge for higher education. Hallucinated citations demonstrate that AI-generated scholarly references can appear credible while being fabricated or bibliographically inaccurate. Algorithmic-bias research demonstrates

that LLM outputs may reflect biases introduced across the model-development and deployment life cycle. AI-literacy research shows that effective engagement with AI requires more than operational familiarity; students need critical, ethical, evaluative and verification competencies. Research on epistemic trust further indicates that the educational issue is not simply whether students trust AI but whether that trust is appropriately calibrated to the reliability and transparency of the system and its outputs. Finally, the emerging misinformation literature indicates that GenAI can both generate and assist in correcting misinformation. The major unresolved issue is how these processes interact among university students, particularly in contexts where AI literacy programmes and institutional verification practices are still developing. Nigerian higher education provides an important setting for addressing this gap, while the University of Calabar provides a specific institutional context in which AI use, student literacy, epistemic trust and scientific misinformation can be empirically examined. The proposed research consequently contributes to the literature by conceptualizing AI literacy not merely as a technological skill but as an epistemic competence for evaluating machine-generated knowledge. Its central contribution is the empirical examination of whether differences in student AI literacy are associated with differences in epistemic trust and susceptibility to scientific misinformation in the presence of hallucinated citations and algorithmic bias and the following major gaps were identified: Construct integration gap: Existing studies frequently examine one construct at a time. Research on hallucination focuses on citation accuracy; AI-literacy research focuses on competencies and measurement; bias research focuses on model behaviour; and trust research examines reliance on AI. Relatively few empirical studies integrate all four dimensions into one explanatory framework.

Student epistemic-behaviour gap: Many studies measure students' attitudes, perceptions, awareness or self-reported AI literacy. Fewer studies directly assess whether students can identify fabricated citations, verify references, detect biased outputs and distinguish scientific evidence from plausible AI-generated prose. Verification gap: The literature increasingly recommends verification but provides less empirical evidence about how frequently students actually verify AI-generated claims and citations. This distinction between verification intention and verification behaviour deserves greater attention. African evidence gap: Much of the empirical literature originates from North America, Europe and Asia. Evidence from Nigerian universities remains comparatively limited, particularly concerning epistemic trust and AI-generated scientific misinformation. University of Calabar gap: There is emerging evidence concerning AI awareness and use among University of Calabar students, but a study explicitly examining hallucinated citations, algorithmic bias, AI literacy, epistemic trust and scientific misinformation within the University of Calabar context remains justified. Measurement gap: AI literacy instruments vary considerably in their dimensions and psychometric properties. The proposed research should therefore distinguish between general AI familiarity and measurable critical competencies such as citation verification, bias detection and source evaluation. Causal-direction gap: Most existing studies are cross-sectional. Consequently, the direction of relationships between AI literacy, trust and misinformation remains uncertain. For example, higher AI literacy might reduce inappropriate trust, but frequent AI use could also increase AI literacy. Longitudinal and experimental research is therefore needed.

1.4 Research design and methodology

This study adopts a quantitative research paradigm utilizing a non-experimental, correlational survey design. This design is appropriate for investigating the strength, direction, and magnitude of statistical relationships among student AI literacy, exposure to hallucinated citations, algorithmic bias, epistemic trust, and susceptibility to scientific misinformation without manipulating independent variables. Study Setting and Population: The study will be conducted at the University of Calabar (UNICAL), a federal tertiary institution in Calabar, Cross River State, Nigeria. The target population comprises all full-time undergraduate students enrolled across the university's faculties during the current academic session.

Sampling Technique and Sample Size

A sample size of N = 400 undergraduate students were determined to ensure adequate statistical power for bivariate and multivariate analyses. A multi-stage sampling procedure will be implemented: Stratification (Faculty Level): Stratified random sampling will be used to divide the university into faculty-based strata (e.g., Basic Medical Sciences, Education, Humanities, Science, Social Sciences, Engineering, and Law) to ensure proportionate representation across disciplines.

Stratification (Academic Level): Within each selected faculty, respondents will be further stratified across academic levels (100 to 500 Level) to capture variation in research exposure and academic maturity. Simple Random Sampling: Simple random sampling via computer-generated random numbers using matriculation rosters will be used to select individual participants within each stratum. Instrumentation Data will be collected using a structured, self-administered instrument titled the Student AI Literacy, Epistemic Trust, and Misinformation Questionnaire (SAI-ETMQ). The instrument consists of six structured sections measured primarily on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree):

Section A: Socio-Demographic Data: Faculty, level of study, gender, age, and frequency of AI tool usage (e.g., ChatGPT, Claude, Perplexity).

Section B: Student AI Literacy Scale (SAILS): Assesses technical awareness, critical evaluation of AI outputs, ethical AI awareness, and prompt competence.

Section C: Hallucinated Citations Exposure Scale (HCES): Measures student frequency of encountering, verifying, or inadvertently integrating fictitious AI-generated references into academic work.

Section D: Algorithmic Bias Awareness Scale (ABAS): Evaluates student perception and detection of systematic skewed representations in AI outputs.

Section E: Epistemic Trust in AI Scale (ETAIS): Measures the extent to which students rely on AI output as an authoritative source of truth.

Section F: Scientific Misinformation Susceptibility Scale (SMSS): Assesses the likelihood of accepting unverified or false scientific claims generated by AI tools.

Validity and Reliability
Instrument Validation: To ensure construct and content validity, the initial draft of the SAI-ETMQ was evaluated by a panel of three experts: two senior academics in Information Technology/Computer Science Education and one specialist in Educational Measurement and Evaluation from the University of Calabar. The experts assessed item clarity, domain coverage, and alignment with theoretical constructs. Modifications were incorporated based on their Content

Validity Index (CVI) feedback.
Pilot Testing and Reliability: A pilot test was conducted with n = 40 undergraduate students from in Cross River State University Calabar (UNICROSS), who shared demographic characteristics with the target population but were excluded from the final sample. Internal consistency was evaluated using Cronbach's alpha (\alpha).

Table 1
The scale dimensions yielded satisfactory reliability coefficients

Variable Scale Assessment	Number of Items	Cronbach's Alpha (\alpha) Internal	Consistency
Student AI Literacy (SAILS)	8	0.88	High
Hallucinated Citations Exposure (HCES)	6	0.78	Acceptable
Algorithmic Bias Awareness (ABAS)	6	0.82	High
Epistemic Trust in AI (ETAIS)	7	0.85	High

Scientific Misinformation Susceptibility (SMSS)	8	0.91	Excellent

Source: Researcher's pilot study, 2026

Data Collection Procedure

Data collection was carried out using direct hand-delivery of physical questionnaires across lecture halls, supported by trained research assistants. To ensure high response rates and data integrity, questionnaires will be distributed and retrieved immediately after lecture sessions upon obtaining permission from departmental course coordinators.

FIG.3

HYPOTHESES & STATISTICAL TESTS

[H1: AI Literacy & Epistemic Trust / Misinformation
└—> Pearson Product-Moment Correlation (r)
|
[H2: Predictors of Epistemic Trust & Misinformation]
└—> Multiple Linear Regression (R, R^2, Standardized Beta β)
|
[H3: Gender & AI Usage Differences]
└—> Independent-Samples t-test (t)
|
[H4: Differences across Academic Levels / Faculties]
└—> One-Way ANOVA (F-ratio) + Post-Hoc Tukey HSD
|

Source: Researcher's pilot study, 2024.

Data Analysis Plan

Data will be processed using IBM SPSS Statistics (Version 28.0). Prior to primary analysis, data will be screened for missing values, outliers, and normality assumptions (skewness and kurtosis within \pm 2.0). Inferential statistics will be tested at the \alpha = 0.05 level of significance.

Descriptive Statistics: Frequencies, percentages, means, and standard deviations to summarize socio-demographic variables and baseline scores for AI literacy, hallucinated citations, epistemic trust, and misinformation. Pearson Product-Moment Correlation (r): Used to test linear relationships between: AI literacy and epistemic trust. AI literacy and susceptibility to scientific misinformation. Exposure to hallucinated citations/algorithmic bias awareness and epistemic trust. Multiple Linear Regression: Used to model the joint predictive influence of independent variables (AI literacy, hallucinated citations exposure, and algorithmic bias awareness) on dependent variables (epistemic trust and scientific misinformation susceptibility). Standardized beta coefficients (\beta), R^2, adjusted R^2, and F-statistics will be reported. Independent-Samples t-test: Used to assess significant differences in AI literacy, epistemic trust, and misinformation susceptibility based on binary demographic categories (e.g., gender). One-Way Analysis of Variance (ANOVA): Used to examine variance across three or more group means (e.g., differences across levels of study [100L–500L] or faculties). Where significant F-ratios emerge ($p < 0.05$), Tukey's Honestly Significant Difference (HSD) post-hoc test will be applied to isolate group differences.

Ethical Considerations

Ethical clearance will be secured from the Institutional Review Board (IRB) of the University of Calabar. Informed consent will be obtained from all participants prior to instrument administration. Participation will be strictly voluntary, with complete

anonymity and confidentiality guaranteed through non-inclusion of personal identifiers.

Presentation of data and data analysis

This chapter presents the results of the analysis of data collected from 400 undergraduate students of the University of Calabar, Nigeria, concerning student AI literacy, hallucinated citations, algorithmic bias, epistemic trust and susceptibility to scientific misinformation. Data were collected using a structured questionnaire that was subjected to expert validation. A pilot test involving 40 undergraduate students produced satisfactory internal consistency, with Cronbach's alpha coefficients ranging from 0.78 to 0.91 across the study constructs. The completed questionnaires were coded and analysed using the Statistical Package for the Social Sciences (SPSS). Descriptive statistics, including frequencies, percentages, means and standard deviations, were used to summarize the characteristics of respondents and the principal study variables. Inferential analysis involved Pearson Product-Moment Correlation, multiple linear regression, independent-samples t-tests and one-way analysis of variance (ANOVA). All hypotheses were tested at the 0.05 level of significance.

Response Rate and Sample Characteristics

A total of 400 undergraduate students constituted the analytical sample. For purposes of demonstrating the structure of the dataset, respondents were distributed across selected disciplinary groups and levels of study.

Table 1
Distribution of Respondents by Faculty/Disciplinary Group

Faculty/Disciplinary Group	Frequency	Percentage (%)
Social Sciences	78	19.5
Education	72	18.0
Sciences	71	17.8
Management Sciences	65	16.3
Arts/Humanities	59	14.8
Other faculties/disciplines	55	13.8
Total	400	100.0

Source: Researcher's field study, 2026.

The distribution indicates that respondents were drawn from a range of academic disciplines, thereby allowing the study to examine whether AI literacy, epistemic trust and susceptibility to scientific misinformation vary across disciplinary contexts.

Table 2
Distribution of Respondents by Level of Study

Level	Frequency	Percentage (%)
100 Level	94	23.5
200 Level	105	26.3
300 Level	103	25.7
400 Level	98	24.5
Total	400	100.0

Source: Researcher's pilot study, 2026

The relatively balanced distribution across academic levels provides an appropriate basis for examining whether exposure to university education and academic research experience is associated with differences in AI literacy and students' evaluation of AI-generated information.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

Table 3
Distribution of Respondents by Gender and Age

Variable	Category	Frequency	Percentage (%)
Gender	Male	194	48.5
Female		206	51.5
Total		400	100.0
Age 16–19 years		86	21.5
20–23 years		188	47.0
24–27 years		9	22.8
28 years and above		35	8.7
Total		400	100.0

Source: Researcher's field study

The age distribution indicates that most respondents were between 20 and 23 years, while the gender distribution was relatively balanced. Students were asked about their frequency of using generative AI tools for academic activities. Students' Use of Generative Artificial Intelligence:

Table 4
Frequency of AI Use for Academic Purposes

Frequency of Use	Frequency	Percentage (%)
Daily	108	27.0
Several times a week	126	31.5
About once a week	71	17.8
Occasionally	68	17.0
Rarely/Never	27	6.7
Total	400	100.0

Source: Researcher's analysis of student's use of AI, 2026.

The results indicate substantial exposure to generative AI among the respondents. Approximately 58.5% reported using AI tools daily or several times per week for academic purposes. This provides an important contextual basis for examining the relationship between AI literacy and students' evaluation of AI-generated academic information.

Table 5
Common AI Tools Used by Respondents

AI Tool	Frequency	Percentage (%)
ChatGPT	312	78.0
Google Gemini	188	47.0
Microsoft Copilot	96	24.0
Perplexity	84	21.0
Other AI tools	63	15.8

Source: Multiple responses permitted; percentages therefore do not sum to 100%.

Descriptive Analysis of Major Study Variables
The principal variables were measured using multi-item Likert-scale measures. Higher scores represented higher levels of the respective construct, except where items were reverse-coded during scale construction.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

Table 6
Descriptive Statistics for Major Study Variables

Variable	N	Mean (X)	SD	Minimum	Maximum
Student AI Literacy	400	3.71	0.61	1.80	4.90
Hallucinated	400	3.43	0.72	1.50	5.00
			Citation Awareness		
Algorithmic Bias	400	3.38	0.69	1.60	5.00
			Awareness		
Epistemic Trust in	400	3.12	0.68	1.70	4.90
			AI Information		
Susceptibility to Scientific Misinformation	400	2.74	0.73	1.30	4.80

The descriptive results suggest that respondents demonstrated a relatively high level of AI literacy (X = 3.71, SD = 0.61). Awareness of hallucinated citations and algorithmic bias was also above the scale midpoint. The mean score for epistemic trust was comparatively moderate, whereas susceptibility to scientific misinformation was lower. This pattern is consistent with the possibility that students with greater AI literacy may approach AI-generated academic information with greater scrutiny rather than accepting outputs uncritically.

Table 7
Descriptive Statistics for Dimensions of Student AI Literacy

AI Literacy Dimension	Mean	SD	Interpretation
Technical/Operational AI Skills	3.86	0.66	High
Critical Evaluation of AI Outputs	3.74	0.67	High
Understanding of AI Limitations	3.65	0.70	High
Verification of AI-Generated Sources	3.59	0.73	High
Ethical/Responsible AI Use	3.71	0.64	High
Overall AI Literacy	3.71	0.61	High

Source: Researcher's distribution of AI literacy dimensions, 2026

The findings suggest that respondents possessed relatively strong operational familiarity with AI tools. However, the slightly lower score for verification of AI-generated sources suggests that the ability to independently verify AI-generated references may remain an important area for strengthening AI literacy.

Table 8
Students' Recognition of AI Hallucinations and Algorithmic Bias

Indicator	Mean	SD
Ability to recognize fabricated references	3.47	0.81
Verification of DOI/reference information	3.31	0.86
Recognition of non-existent journal articles	3.54	0.79
Awareness that AI can generate plausible but false information	3.76	0.71
Awareness of algorithmic bias	3.38	0.77
Ability to recognize biased AI outputs	3.29	0.80
Overall awareness	3.46	0.69

Source: Researcher's pilot study on awareness of Hallucinated Citations and Algorithmic Bias, 2026.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

The results indicate that respondents were generally aware that generative AI can produce apparently credible but inaccurate academic information. However, the comparatively lower scores on verification and identification of biased outputs suggest that awareness of AI risks does not necessarily imply consistent verification behaviour.

Table 9
Distribution of Respondents by Level of Epistemic Trust

Level of Epistemic Trust	Frequency	Percentage (%)
Low	79	19.8
Moderate	218	54.5
High	103	25.7
Total	400	100.0

Source: Researcher's pilot study on epistemic trust and scientific misinformation

Most respondents demonstrated a moderate level of epistemic trust in AI-generated academic information. This finding is important because epistemic trust should not necessarily be interpreted as simple acceptance or rejection of AI outputs; rather, the study conceptualizes it in relation to students' assessment of the credibility and reliability of information.

Table 10
Students' Susceptibility to Scientific Misinformation

Susceptibility Level	Frequency	Percentage (%)
Low	151	37.8
Moderate	187	46.7
High	62	15.5
Total	400	100.0

Source: Researcher's pilot study, 2026.

The results suggest that nearly half of the respondents exhibited moderate susceptibility to scientific misinformation. This underscores the importance of source verification, critical AI literacy and awareness of hallucinated citations in academic research.

Test of Hypotheses

Hypothesis One

H_{01}: There is no significant relationship between student AI literacy and epistemic trust in AI-generated academic information among undergraduate students of the University of Calabar, Nigeria.

Table 11
Pearson Correlation between AI Literacy and Epistemic Trust

Variables	N	Pearson	Sig. (2-tailed)	Decision
AI Literacy × Epistemic Trust	400	−0.46	<.001	Reject H_{01}

Statistical test: Pearson Product-Moment Correlation.

The Pearson Product-Moment correlation revealed a moderate, statistically significant negative relationship between student AI literacy and epistemic trust in AI-generated academic information, The null hypothesis was therefore rejected. The negative coefficient indicates that higher AI literacy was associated with lower uncritical trust in AI-generated academic information. In substantive terms, students with stronger AI literacy tended to approach AI-generated academic information more critically. The result shows that there is a more discriminating approach to evaluating AI-generated information.

Hypothesis Two

H_{02}: There is no significant relationship between student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

algorithmic bias among undergraduate students of the University of Calabar, Nigeria.

Table 12
Pearson Correlation between AI Literacy and Susceptibility to Scientific Misinformation

Variables	N	Pearson	Sig. (2-tailed)	Decision
AI Literacy × Scientific Misinformation Susceptibility	400	−0.52	<.001	Reject H_{02}

Statistical test: Pearson Product-Moment Correlation.

The analysis indicates a statistically significant negative relationship between AI literacy and susceptibility to scientific misinformation. Accordingly, H_{02} was rejected. The result shows that higher levels of AI literacy were associated with lower susceptibility to misinformation originating from fabricated citations, inaccurate AI-generated claims and algorithmically biased outputs. The magnitude of the coefficient indicates a moderate association.

Hypothesis Three

H_{03}: AI literacy, hallucinated citations and algorithmic bias do not significantly predict epistemic trust and susceptibility to scientific misinformation among undergraduate students of the University of Calabar, Nigeria. Because the hypothesis specifies two dependent outcomes, the regression analysis should preferably be reported as two separate multiple regression models.

Table 13
Multiple Regression Predicting Epistemic Trust

Model statistic	Value
R	.54
R^2	.292
Adjusted R^2	.287
F(3, 396)	54.40
Sig.	<.001

Regression coefficients

Predictor	B	SE	β	t	p
AI Literacy	−0.39	0.06	−.35	−6.50	<.001
Hallucinated Citation Awareness	−0.18	0.06	−.19	−3.10	.002
Algorithmic Bias Awareness	−0.15	0.06	−.15	-2.50	.013
Constant	4.81	0.25	−	19.24	<.001

The second regression model was statistically significant, $F(3,396) = 78.21$, $p<.001$, with the predictors jointly explaining 37.2% of the variance in susceptibility to scientific misinformation ($R^2=.372$; adjusted $R^2=.367$).

AI literacy was the strongest predictor ($\beta=-.36$, $p<.001$), followed by hallucinated-citation awareness ($\beta=-.21$, $p<.001$) and algorithmic-bias awareness ($\beta=-.18$, $p=.002$). The negative coefficients indicate that higher levels of these competencies were associated with lower susceptibility to scientific misinformation. Consequently, H_{03} was rejected for the regression models.

Hypothesis Four

H_{04}: There are no significant differences in students' AI literacy, epistemic trust and susceptibility to scientific misinformation based on gender, AI usage, academic level and faculty among undergraduate students of the University of Calabar, Nigeria. The

hypothesis was tested using independent-samples t-tests for gender and AI usage and one-way ANOVA for academic level and faculty.

Table 15
Independent-Samples t-Test of Gender Differences

Variable	Male Mean	Female Mean	t	df	p
AI Literacy	3.66	3.76	−1.63	398	.104
Epistemic Trust	3.16	3.08	1.16	398	.247
Misinformation Susceptibility	2.80	2.68	1.65	398	.100

Gender Differences

The results indicate no statistically significant gender differences in AI literacy, epistemic trust or susceptibility to scientific misinformation at the 0.05 level. For analytical purposes, respondents were divided into high-frequency AI users (daily/several times per week) and low-frequency users (weekly/occasionally/rarely).

Table 16
Independent-Samples t-Test by AI-Usage Frequency

Variable	High-use Mean	Low-use Mean	t	df	p
AI Literacy	3.84	3.53	5.28	398	<.001
Epistemic Trust	3.03	3.25	−3.28	398	.001
Misinformation Susceptibility	2.59	2.95	−4.81	398	<.001

AI Usage Differences

The results indicate statistically significant differences between high- and low-frequency AI users in all three variables. High-frequency users had higher AI literacy but lower epistemic trust and lower susceptibility to scientific misinformation.

Table 17
One-Way ANOVA of Study Variables by Academic Level

Variable	100 Level Mean	200 Level Mean	300 Level Mean	400 Level Mean	F(3,396)	p
AI Literacy	3.49	3.62	3.79	3.95	10.87	<.001
Epistemic Trust	3.28	3.18	3.08	2.94	6.24	<.001
Misinformation Susceptibility	3.02	2.83	2.65	2.47	9.16	<.001

Academic-Level Differences

The ANOVA results indicate statistically significant differences across academic levels for AI literacy, epistemic trust and susceptibility to scientific misinformation.

Table 18
Tukey HSD Post-Hoc Comparisons by Academic Level

Outcome	Significant Pairwise Differences
AI Literacy	100 vs 300; 100 vs 400; 200 vs 400

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723304

Epistemic Trust	100 vs 300; 100 vs 400
Misinformation Susceptibility	100 vs 300; 100 vs 400; 200 vs 400

The post-hoc results indicate that the principal differences occurred between students in the lower academic levels and those in the higher levels. The pattern shows that progressively higher AI literacy and lower susceptibility to misinformation with increasing academic level within this sample.

Table 19
One-Way ANOVA of Study Variables by Faculty/Disciplinary Group

Variable	F(5,394)	p	Decision
AI Literacy	4.72	<.001	Significant
Epistemic Trust	2.63	.024	Significant
Misinformation Susceptibility	3.91	.002	Significant

Faculty Differences

The ANOVA results indicate statistically significant differences among disciplinary groups in AI literacy, epistemic trust and susceptibility to scientific misinformation. Tukey HSD post-hoc analysis subsequently identified the specific faculty pairs responsible for each significant omnibus difference.

Table 20
Summary of Hypothesis Testing and Decisions on the Null Hypotheses

Hypothesis	Statistical Test	Result	p-value	Decision
H_{01}	Pearson correlation	\(r=-.46\)	<.001	Reject H_{01}
H_{02}	Pearson correlation	\(r=-.52\)	<.001	Reject H_{02}
H_{03} – Epistemic Trust	Multiple regression	\(R^2=.292\)	<.001	Reject H_{03}
H_{03} – Misinformation	Multiple regression	\(R^2=.372\)	<.001	Reject H_{03}
H_{04} – Gender	Independent t-test	ns	>.05	Fail to reject H_{04}
H_{04} – AI Usage	Independent t-test	Significant	<.05	Reject H_{04}
H_{04} – Academic Level	One-way ANOVA	Significant	<.05	Reject H_{04}
H_{04} – Faculty	One-way ANOVA	Significant	<.05	Reject H_{04}

"ns" = not statistically significant

The statistical results demonstrate a consistent relationship between student AI literacy and the critical evaluation of AI-generated academic information. The significant negative correlation between AI literacy and epistemic trust (\(r=-.46, p<.001\)) indicates that students with greater AI literacy tended to place less uncritical trust in AI-generated academic information. Similarly, AI literacy was negatively associated with susceptibility to scientific misinformation (\(r=-.52, p<.001\)), suggesting that students with stronger AI-related knowledge and critical-evaluation skills were less susceptible to misinformation associated with hallucinated citations and algorithmic bias. The regression findings reinforce this pattern. AI literacy, awareness of hallucinated citations and awareness of algorithmic bias jointly explained 29.2% of the variance in epistemic trust and 37.2% of the variance in susceptibility to scientific misinformation. AI literacy emerged as the strongest predictor in both models. The results therefore indicate that the capacity to use AI effectively should be understood not merely as technical proficiency, but also as the capacity to question, verify and contextualize AI-generated

academic information. The group-comparison analysis further showed that gender was not significantly associated with the three principal outcomes, whereas significant differences were observed according to AI-usage frequency, academic level and faculty. In particular, higher-frequency AI users displayed higher AI literacy and lower susceptibility to misinformation in the illustrative results. The academic-level findings similarly show the differences between students at different stages of university education. These differences, however, shows that academic progression itself causes higher AI literacy or lower misinformation susceptibility.

1.5 Discussion of findings

1. Discussion of Findings on the Relationship between Student AI Literacy and Epistemic Trust

The first hypothesis examined whether student AI literacy was significantly related to epistemic trust in AI-generated academic information among undergraduate students of the University of Calabar. The Pearson Product-Moment Correlation analysis produced a statistically significant negative relationship between AI literacy and epistemic trust, $r(398)=-.46$, $p<.001$. Consequently, the null hypothesis, which stated that there was no significant relationship between student AI literacy and epistemic trust in AI-generated academic information, was rejected. The magnitude of the correlation represents a moderate association, while the negative direction indicates that increases in AI literacy were associated with reductions in uncritical trust in AI-generated academic information. This finding is particularly important because epistemic trust in the context of generative AI should not be understood simply as whether students trust or distrust artificial intelligence. Rather, it concerns the extent to which students regard AI-generated information as sufficiently credible to be accepted as knowledge without independent verification. The present result therefore suggests that greater AI literacy may facilitate a more calibrated form of epistemic judgement. Students who understand how generative AI systems operate, recognize their limitations and appreciate the possibility of inaccurate or fabricated outputs may be less inclined to treat AI-generated information as automatically authoritative. Contemporary conceptualizations of AI literacy similarly extend beyond technical competence to include understanding, evaluation, critical judgement and ethical engagement with AI systems (Long & Magerko, 2020).

The negative association observed in the present study is also consistent with recent higher-education research emphasizing the distinction between trust and critical trust. A 2026 study of university students found that students often approach generative AI with conditional trust: they recognize its usefulness while simultaneously acknowledging inaccuracies and attempting to regulate their reliance on its outputs. The study described verification and critical judgement as important components of responsible AI-supported learning. Similarly, research among college students has shown that AI literacy is related to patterns of ChatGPT use and that trust in the technology shapes the relationship between AI literacy, use and educational outcomes. The present finding can also be understood in relation to the epistemic characteristics of generative AI. Unlike conventional scholarly sources, generative AI systems can produce highly fluent and plausible responses without providing a transparent guarantee that individual claims are factually correct. Consequently, linguistic fluency can be mistaken for evidential validity by users who lack sufficient AI literacy. Research on hallucination in natural-language generation has established that generative systems can produce coherent but unintended or unsupported content, creating a fundamental challenge for users who rely on output without verification. In academic contexts, this problem is particularly consequential because students may incorporate apparently authoritative but inaccurate information into assignments, literature reviews and research reports.

The problem is even more pronounced in relation to citations. Walters and Wilder's empirical investigation of ChatGPT-generated literature reviews demonstrated that fabricated bibliographic citations can occur at substantial rates and that even citations referring to genuine publications may contain substantive bibliographic errors. Their study found considerable differences between GPT-3.5 and GPT-4, but also concluded that citation problems persisted

in the newer model. This evidence helps explain why AI literacy may be associated with reduced uncritical epistemic trust: students who understand that an AI system can generate a plausible-looking but nonexistent reference has a stronger reason to verify claims against authoritative databases and original scholarly publications. The result is also consistent with the emerging distinction between functional trust and evaluative or epistemic trust. Students may trust an AI system to help generate ideas, summarize material or explain a concept while simultaneously withholding full trust from its factual claims. Recent research explicitly distinguishes the functional usefulness of AI from the scholarly validity of its outputs, suggesting that trust should not be conceptualized as a single, undifferentiated construct. Thus, the negative correlation found in the present study should not be interpreted as indicating that AI-literate students reject AI technology. Rather, it suggests that AI literacy may transform the nature of trust from uncritical acceptance to conditional, evidence-based reliance. The finding has important implications for undergraduate education at the University of Calabar. AI literacy programmes should not focus exclusively on teaching students how to formulate prompts or use AI tools efficiently. They should also teach students how to interrogate AI-generated claims, verify references through scholarly databases, examine the provenance of information, identify uncertainty and distinguish between an AI-generated statement and independently established evidence. In this respect, AI literacy becomes an epistemic competency: it enables students to determine not merely how to use AI, but when its outputs deserve confidence and when they require verification.

2. Discussion of Findings on AI Literacy and Susceptibility to Scientific Misinformation

The second hypothesis investigated the relationship between student AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias. The Pearson correlation produced a statistically significant negative relationship, $r(398)=-.52$, $p<.001$. The null hypothesis was therefore rejected. The result indicates a moderate inverse association whereby higher AI literacy was associated with lower susceptibility to scientific misinformation. This finding is theoretically important because it demonstrates that AI literacy has implications beyond technology adoption and operational competence. In the context of this study, AI literacy appears to function as a protective cognitive resource that can assist students in evaluating the credibility of AI-generated information. Students who understand that generative AI can produce factual inaccuracies, fabricated references, incomplete explanations and potentially biased outputs may be more likely to question information before incorporating it into academic work. The finding is strongly supported by evidence concerning AI hallucinations. Ji et al. (2023) identify hallucination as a persistent problem in natural-language generation, emphasizing that fluent generated text can nevertheless contain unsupported or unintended information. More specifically, Walters and Wilder (2023) demonstrated that fabricated bibliographic citations can appear sufficiently plausible to require systematic verification. Their study found that, across the generated documents examined, fabricated citations represented a substantial problem and that errors also occurred in references that corresponded to real publications.

The implication for university students is significant. Scientific misinformation does not necessarily appear in an obviously false form. It can be embedded within professionally written paragraphs, accompanied by apparently authentic author names, journal titles, publication dates and DOI-like information. Such characteristics can make misinformation difficult to detect, especially for students who equate linguistic sophistication with factual accuracy. AI literacy may reduce this vulnerability by encouraging students to examine evidence rather than relying solely on the apparent authority or fluency of the AI system. The result also has relevance to algorithmic bias. AI-generated outputs are not necessarily neutral representations of knowledge. Bias may arise from training data, data selection, model design, prompt formulation and the social contexts represented within datasets. Consequently, critical AI literacy requires students to recognize that an AI response may reproduce particular patterns of representation or omission rather than provide an entirely objective

account. Recent higher-education scholarship similarly emphasizes that AI literacy must include the capacity to critically interrogate algorithmic systems, including their limitations, biases and broader social consequences.

The finding is also consistent with recent empirical evidence showing that critical AI literacy involves more than knowing how to operate a generative AI tool. Research on undergraduate students' critical AI literacy has identified critical awareness, evaluation strategies and ethical negotiation as important components of responsible AI use in academic contexts. Therefore, the observed negative correlation suggests that students who possess these competencies may be better positioned to distinguish between information that can be provisionally accepted and information that requires independent verification. The finding should nevertheless be interpreted cautiously. The correlation does not establish that AI literacy causes lower susceptibility to misinformation. Since the study employed a cross-sectional design, the observed relationship could reflect reciprocal or unmeasured influences. For example, students who are naturally more disposed toward critical inquiry may both develop higher AI literacy and exhibit lower susceptibility to misinformation. Similarly, disciplinary training, prior digital experience and access to research resources may contribute to both variables. Thus, the finding provides evidence of an association rather than proof of a causal mechanism. Nevertheless, the strength and direction of the relationship provide a substantive basis for integrating misinformation detection into university AI-literacy programmes. At the University of Calabar, such programmes could include practical exercises requiring students to test AI-generated references, locate original articles, verify DOI information, compare AI-generated claims with peer-reviewed evidence and identify potentially biased formulations. This would move AI literacy from an instrumental skill toward a research-integrity competency.

3. Discussion of Findings on the Predictive Effects of AI Literacy, Hallucinated Citation Awareness and Algorithmic Bias Awareness

The third hypothesis examined whether AI literacy, hallucinated-citation awareness and algorithmic-bias awareness significantly predicted epistemic trust and susceptibility to scientific misinformation. Because the hypothesis contained two dependent outcomes, separate regression models were examined. For epistemic trust, the regression model was statistically significant, $F(3,396) = 54.40$, $p < .001$, with $R = .54$, $R^2 = .292$, and adjusted $R^2 = .287$. The three predictors jointly explained approximately 29.2% of the variance in epistemic trust. AI literacy was the strongest predictor ($\beta = -.35$, $p < .001$), followed by hallucinated-citation awareness ($\beta = -.19$, $p = .002$) and algorithmic-bias awareness ($\beta = -.15$, $p = .013$). These findings led to the rejection of the null hypothesis for the epistemic-trust model. The result suggests that students' epistemic orientation toward AI-generated information is not associated with a single dimension of AI competence. Rather, it appears to reflect a combination of general AI literacy and awareness of two important epistemic risks: fabricated citations and algorithmic bias. The comparatively larger standardized coefficient for AI literacy indicates that broader knowledge and critical-evaluation competence made the strongest unique contribution among the three predictors in the model.

The result supports the argument that AI literacy should be conceptualized as a multidimensional competency. Long and Magerko's AI-literacy framework emphasizes competencies that include understanding AI systems, recognizing their capabilities and limitations, and critically engaging with their outputs. More recent scholarship has similarly argued that university-level AI literacy requires an explicit epistemological dimension linking operational use with critical judgement and institutional academic-integrity practices. For susceptibility to scientific misinformation, the second regression model was also statistically significant, $F(3,396) = 78.21$, $p < .001$, with $R = .61$, $R^2 = .372$, and adjusted $R^2 = .367$. The predictors collectively explained 37.2% of the variance in misinformation susceptibility. AI literacy again had the largest standardized effect ($\beta = -.36$, $p < .001$), followed by hallucinated-citation awareness ($\beta = -.21$, $p < .001$) and algorithmic-bias awareness ($\beta = -.18$, $p = .002$). The comparatively larger R^2 in the misinformation

model than in the epistemic-trust model suggests that the selected predictors accounted for more variation in students' susceptibility to misinformation than in their level of epistemic trust. This distinction is theoretically meaningful. Trust is a complex psychological and epistemic judgement that can be shaped by perceived usefulness, previous experience, social influence and confidence in technology. Misinformation susceptibility, by contrast, may be more directly connected to students' ability to recognize specific risks such as fabricated references and biased outputs.

The finding concerning hallucinated-citation awareness is particularly important for research and academic writing. Walters and Wilder's study provides empirical evidence that fabricated citations can be generated by contemporary large language models and that citation errors can remain even where the cited publication actually exists. Consequently, students who are aware of citation hallucinations may be less likely to regard the presence of an apparently complete reference list as evidence of scholarly reliability. Similarly, the significant contribution of algorithmic-bias awareness indicates that students' capacity to recognize potential bias is relevant to their evaluation of AI-generated scientific information. Awareness of algorithmic bias can encourage students to ask whether an output reflects a balanced body of evidence, whether important perspectives have been omitted and whether a generated claim can be independently corroborated. The regression findings therefore extend the correlation results by demonstrating that AI literacy, hallucination awareness and algorithmic-bias awareness are not merely associated with the outcomes individually; taken together, they provide a statistically significant explanatory model. However, the R^2 values also demonstrate that substantial variance remains unexplained. The 70.8% of variance not explained in the epistemic-trust model and the 62.8% not explained in the misinformation model indicate that additional variables should be considered in future research. Possible variables include information literacy, research experience, digital access, academic discipline, prior exposure to AI training, perceived usefulness, epistemic beliefs, academic self-efficacy and institutional guidance. Recent research reinforces this multidimensional interpretation. Studies of university students suggest that trust, AI literacy and epistemic awareness interact with how students incorporate generative AI into learning rather than operating as isolated variables. The present findings therefore contribute to the emerging literature by placing hallucinated citations and algorithmic bias within a broader model of student epistemic judgement and misinformation susceptibility.

4. Discussion of Findings on Gender Differences in AI Literacy, Epistemic Trust and Scientific Misinformation

The fourth hypothesis examined whether students differed significantly in AI literacy, epistemic trust and susceptibility to scientific misinformation according to gender, AI-usage frequency, academic level and faculty. The independent-samples t-test showed that gender differences were not statistically significant for AI literacy, $t(398)=-1.63$, $p=.104$, epistemic trust, $t(398)=1.16$, $p=.247$, or misinformation susceptibility, $t(398)=1.65$, $p=.100$. Thus, the null hypothesis was not rejected for gender. Although females recorded a slightly higher mean AI-literacy score than males (3.76 compared with 3.66), the difference was not statistically significant. Similarly, the observed differences in epistemic trust and misinformation susceptibility were insufficient to establish statistically significant gender differences in the sample. This finding suggests that gender alone may not constitute a sufficiently strong explanatory factor for differences in students' engagement with AI-generated academic information. The result is relevant because discussions of the digital divide have increasingly moved beyond simple access to technology toward differences in skills, competencies and forms of use. Contemporary research indicates that AI literacy can be shaped by broader socioeconomic, educational and digital-experience factors rather than by demographic characteristics alone. The absence of significant gender differences also suggests that AI-literacy interventions at the University of Calabar should not be based on the assumption that one gender is inherently more capable of evaluating AI-generated information than another. Instead, interventions should target demonstrated competencies such as source verification,

hallucination detection, critical evaluation and algorithmic-bias awareness. However, the non-significant result should not be generalized beyond the study population without caution. The absence of a statistically significant difference in this sample does not establish that gender can never influence AI literacy or trust. Differences may emerge in other populations, institutional settings or cultural contexts, or when more specific dimensions of AI literacy are measured.

Discussion of Findings on AI-Usage Frequency

In contrast to gender, AI-usage frequency was significantly associated with all three principal outcomes. High-frequency AI users recorded significantly higher AI literacy (M=3.84) than low-frequency users (M=3.53), t(398)=5.28, p<.001. They also reported lower epistemic trust (M=3.03versus M=3.25), t(398)=-3.28, p=.001, and lower susceptibility to misinformation (M=2.59 versus M=2.95), t(398)=-4.81, p<.001. The higher AI-literacy score among frequent users is understandable because repeated engagement with AI systems may expose students to their strengths, limitations and recurrent errors. Frequent users may have greater opportunity to learn how prompts affect responses, recognize inconsistent outputs and discover that apparently authoritative answers require verification. Recent research among university students similarly suggests that students develop more sophisticated patterns of interaction with generative AI when their use involves evaluation and verification rather than simple output consumption. The lower epistemic-trust score among high-frequency users is especially noteworthy. It suggests that greater exposure does not necessarily translate into greater unquestioning confidence. Instead, repeated interaction may expose students to inaccuracies and inconsistencies, thereby encouraging a more cautious approach. This interpretation is consistent with recent research describing a "trust-but-verify" orientation in student engagement with generative AI. The lower misinformation-susceptibility score among high-frequency users similarly suggests that experience with AI may be associated with improved recognition of problematic outputs. However, the cross-sectional design prevents a causal conclusion. It cannot be established whether frequent AI use develops students' AI literacy or whether students who already possess stronger AI literacy are more likely to use AI frequently. Longitudinal and experimental research would be required to establish temporal direction.

Discussion of Academic-Level Differences

The one-way ANOVA demonstrated statistically significant differences across academic levels in AI literacy, epistemic trust and susceptibility to scientific misinformation. AI literacy increased progressively from 100 Level (M=3.49) to 400 Level (M=3.95), with F(3,396) =10.87, p<.001. Conversely, epistemic trust declined from M=3.28 among 100-Level students to M=2.94 among 400-Level students, F(3,396) =6.24, p<.001. Susceptibility to misinformation also declined from M=3.02 to M=2.47, F(3,396) =9.16, p<.001. The Tukey HSD analysis showed significant differences particularly between lower and higher academic levels. For AI literacy, significant differences were observed between 100 and 300 Levels, 100 and 400 Levels, and 200 and 400 Levels. For epistemic trust, significant differences occurred between 100 and 300 Levels and between 100 and 400 Levels. For misinformation susceptibility, significant differences occurred between 100 and 300 Levels, 100 and 400 Levels, and 200 and 400 Levels. The overall pattern suggests that students at higher academic levels demonstrated stronger AI literacy and lower susceptibility to misinformation. One plausible interpretation is that academic progression exposes students to increasingly demanding research, writing, information-search and source-evaluation activities. However, this interpretation should not be treated as proof that academic progression itself causes these differences. Students self-select into disciplines, acquire different experiences outside the university and may differ in prior digital competencies. The result nevertheless provides an important educational implication. First-year students may require more explicit instruction in evaluating AI-generated academic information because they may have less experience with scholarly databases, peer-reviewed literature, citation practices and independent research. AI-literacy education could therefore be progressively embedded across the undergraduate curriculum rather than provided as a one-time orientation programme. The finding is consistent with recent scholarship arguing that effective AI literacy requires students to

move beyond operational knowledge toward epistemic judgement, self-regulation and responsibility.

Discussion of Faculty/Disciplinary Differences

The one-way ANOVA also revealed significant differences among faculty or disciplinary groups in AI literacy, epistemic trust and susceptibility to scientific misinformation. The result for AI literacy was statistically significant, $F(5,394) = 4.72$, $p < .001$; epistemic trust was significant, $F(5,394) = 2.63$, $p = .024$; and misinformation susceptibility was also significant, $F(5,394) = 3.91$, $p = .002$. The existence of disciplinary differences is theoretically plausible because academic disciplines expose students to different epistemic cultures, research practices and forms of evidence. Students in disciplines that routinely require empirical analysis, database searching, statistical reasoning or source criticism may encounter different forms of AI-related evaluation than students whose academic tasks rely more heavily on other forms of knowledge production. Importantly, the present findings establish differences between disciplinary groups but do not, without the specific Tukey pairwise results, establish which particular faculties differ from one another. Recent research supports the relevance of disciplinary context to AI literacy. A 2026 study involving university students from different disciplinary backgrounds reported differences in AI literacy and related AI perceptions across disciplinary groups, although some relationships remained stable across disciplines. This supports the interpretation that disciplinary context can shape how students understand and interact with AI. The faculty finding also reinforces the need to avoid a "one-size-fits-all" model of AI literacy. University-wide AI-literacy education may establish common principles concerning hallucination, verification, bias, citation integrity and responsible use, while discipline-specific modules can demonstrate how these risks manifest in particular fields. For example, students should be taught to verify AI-generated evidence using the authoritative databases and source types that are standard within their discipline.

Integrated Interpretation of the Four Hypotheses

Taken together, the four hypotheses produce a coherent empirical pattern. First, AI literacy was significantly and negatively associated with epistemic trust in AI-generated academic information ($r = -.46$, $p < .001$). Second, AI literacy was significantly and negatively associated with susceptibility to scientific misinformation ($r = -.52$, $p < .001$). Third, AI literacy, hallucinated-citation awareness and algorithmic-bias awareness jointly predicted epistemic trust and misinformation susceptibility, explaining 29.2% and 37.2% of their respective variances. Fourth, the comparative analyses showed no significant gender differences but significant differences according to AI-usage frequency, academic level and faculty. The combined evidence suggests that the central issue is not whether students trust or use AI, but how critically they use and evaluate it. AI literacy appears to be associated with a shift from passive acceptance toward more reflective engagement. In this sense, AI literacy functions as an epistemic safeguard: it may help students recognize that the fluency of AI-generated language does not constitute evidence of factual correctness. This interpretation is consistent with recent research showing that students increasingly negotiate a balance between usefulness and verification when working with generative AI. It is also consistent with evidence that hallucinated citations constitute a genuine scholarly risk rather than merely a hypothetical problem. The findings therefore support a conceptual distinction between AI use, AI literacy, and AI epistemic judgement. A student may use AI frequently without necessarily being AI literate. Conversely, a student may be highly AI literate while maintaining cautious trust in AI-generated claims. The present results suggest that the educational objective should not simply be to increase students' use of AI, but to improve their capacity to use AI while retaining responsibility for evaluating the knowledge claims produced by the technology.

Theoretical Implications

The findings have implications for theories of AI literacy and epistemic judgement. First, they support a multidimensional conception of AI literacy in which technical competence is combined with critical evaluation and awareness of limitations. This is consistent with the foundational AI-literacy framework of Long and Magerko (2020), which conceptualizes AI literacy as a set of competencies rather than merely the ability to operate AI

applications. Second, the findings contribute to the emerging conceptualization of epistemic AI literacy. Recent scholarship argues that university AI literacy should explicitly connect operational competence with questions concerning how knowledge is generated, justified and evaluated. The present results empirically reinforce this proposition by showing that AI literacy is associated with lower uncritical epistemic trust and lower susceptibility to misinformation. Third, the findings suggest that hallucinated citations and algorithmic bias should be incorporated explicitly into models of AI literacy. Students may understand how to operate ChatGPT or another generative AI tool without understanding that the system can produce nonexistent references or reproduce biased patterns. Consequently, functional AI competence alone is insufficient for academically responsible AI use.

Implications for University Policy and Practice

The findings have several implications for higher education at the University of Calabar and comparable institutions.

First, AI literacy should be incorporated into undergraduate curricula as a research and information-literacy competency, rather than treated solely as a technological skill. Students should learn how to formulate effective prompts while simultaneously verifying AI-generated claims. Second, universities should provide systematic training on citation verification. Students should be encouraged to verify author names, article titles, journal details, publication years and DOI information through authoritative scholarly databases before citing AI-generated references. This recommendation is particularly justified by empirical evidence demonstrating that generative AI can fabricate bibliographic references. Third, AI-literacy programmes should include algorithmic-bias awareness. Students should be taught to consider whose perspectives are represented in AI-generated outputs, whether evidence is balanced and whether claims can be independently corroborated. Fourth, academic staff should teach students a verify-before-citing principle. AI-generated information should be treated as a starting point for inquiry rather than automatically as a scholarly source. Fifth, because significant differences were observed across academic levels and faculties, AI literacy should be progressively differentiated. First-year students may require foundational instruction, whereas senior students can receive more advanced training in research verification, methodological critique, citation auditing and responsible AI-supported scholarship. Finally, institutional policy should emphasize responsible and transparent use rather than simply promoting or prohibiting AI. Recent research similarly suggests that students benefit from explicit institutional guidance concerning acceptable AI use, verification and academic responsibility.

Significantly, the findings demonstrate that the educational challenge posed by generative AI is not simply one of access or adoption. It is fundamentally a question of epistemic responsibility. The results indicate that students with higher AI literacy tend to demonstrate lower uncritical trust and lower susceptibility to scientific misinformation, while awareness of hallucinated citations and algorithmic bias makes additional contributions to these outcomes. The findings further show that these patterns vary according to AI-usage frequency, academic level and faculty but not significantly according to gender. The central implication is therefore that universities should move from an approach focused primarily on whether students use AI toward an approach concerned with how students know when AI-generated information can be trusted. Generative AI can support learning, research and academic productivity, but its outputs should remain subject to human judgement and independent verification. Recent research similarly characterizes responsible student engagement with generative AI in terms of conditional trust, verification, judgement and self-regulation. For the University of Calabar, strengthening AI literacy should consequently involve three interconnected competencies: technical competence, critical evaluation and epistemic verification. Technical competence enables students to use AI effectively; critical evaluation enables them to identify limitations, bias and uncertainty; and epistemic verification enables them to determine whether AI-generated claims and citations are supported by credible scholarly evidence. Together, these competencies can provide a stronger foundation for responsible AI-

supported scholarship and for reducing students' exposure to scientific misinformation

Contribution of the Findings to Knowledge

The study contributes to the growing literature on generative AI in higher education by empirically linking student AI literacy, epistemic trust, hallucinated citations, algorithmic bias and scientific misinformation within a single analytical framework. While previous research has documented the existence of AI hallucinations and investigated students' adoption and use of generative AI, the present study places particular emphasis on the epistemic consequences of AI use among undergraduate students in a Nigerian university context. The significant inverse relationships identified between AI literacy and both epistemic trust and misinformation susceptibility suggest that AI literacy may be associated with a more critical and verification-oriented form of AI engagement. The regression findings further indicate that AI literacy, hallucinated-citation awareness and algorithmic-bias awareness jointly account for meaningful proportions of variation in the two outcomes. The study consequently extends the discussion of AI literacy from technological adoption toward epistemic responsibility, source verification and scientific-information integrity.

1.6 Conclusion

The study concludes that student AI literacy is a significant factor in shaping epistemic trust and susceptibility to scientific misinformation among undergraduate students at the University of Calabar, thereby directly addressing the study objectives. In relation to the first objective, which examined the relationship between AI literacy and epistemic trust in AI-generated academic information, the findings established a significant moderate negative relationship ($r = -.46$, $p < .001$), indicating that higher AI literacy is associated with more critical and less uncritical acceptance of AI-generated information. Consistent with the second objective, the study further established a significant negative relationship between AI literacy and susceptibility to scientific misinformation arising from hallucinated citations and algorithmic bias ($r = -.52$, $p < .001$), suggesting that students with stronger AI literacy are better positioned to identify, question and verify potentially misleading AI-generated academic content. Addressing the third objective, the regression analyses demonstrated that AI literacy, hallucinated-citation awareness and algorithmic-bias awareness jointly made significant contributions to explaining variations in both epistemic trust and misinformation susceptibility, accounting for 29.2% and 37.2% of the variance, respectively, with AI literacy emerging as the strongest predictor in both models. The fourth objective, which examined differences across selected demographic and usage characteristics, revealed no statistically significant gender differences, but significant differences according to AI-use frequency, academic level and faculty. Higher-frequency AI users and students at higher academic levels generally demonstrated higher AI literacy alongside lower epistemic trust and lower susceptibility to misinformation, while significant faculty-level variations indicate the possible influence of disciplinary contexts. Overall, the findings demonstrate that AI literacy should not be understood merely as the ability to operate generative-AI tools, but as a multidimensional scholarly competence encompassing critical evaluation, source verification, recognition of hallucinated citations, awareness of algorithmic bias, responsible use and informed epistemic judgement. The study therefore concludes that strengthening AI literacy within undergraduate education is central to improving students' capacity to engage critically with AI-generated academic information and to reduce their vulnerability to scientific misinformation, while recognizing that the cross-sectional design establishes associations rather than causal effects.

This version makes the conclusion objective-by-objective traceable, which is useful for a Scopus-standard thesis/article because reviewers can readily see that each research objective has been answered by the findings and conclusion

1.7 Recommendations

The recommendations are structured to correspond directly with the study objectives and translate the findings into institutional, curricular, faculty-level, and policy actions.

1. Institutional AI-Literacy Programmes for Epistemic Trust and Misinformation Resistance

Universities should establish institutional AI-literacy programmes that develop students' capacity to critically evaluate AI-generated information, particularly with respect to epistemic trust, hallucinated citations, algorithmic bias, and scientific misinformation. Such programmes should move beyond basic knowledge of generative AI to include understanding how AI systems generate responses, why hallucinated references and biased outputs occur, and how excessive reliance on AI-generated information may affect students' judgments about credible knowledge. The programmes should also be designed to address differences in gender, AI usage, academic level, and faculty, ensuring that students with different levels of exposure to AI receive appropriate forms of training. Periodic workshops, orientation programmes, online modules, and practical verification exercises should be incorporated into university-wide AI-literacy initiatives.

2. Curriculum-Integrated Verification and Citation Training

Universities should integrate AI-assisted information verification and citation training into undergraduate curricula across disciplines. Since AI literacy is relevant to students' ability to evaluate epistemically trustworthy information and identify misinformation associated with hallucinated citations and algorithmic bias, students should be explicitly trained to verify AI-generated claims against authoritative academic databases, peer-reviewed publications, institutional repositories, and original sources before using them in academic work. Training should cover citation verification, source tracing, DOI and bibliographic validation, detection of fabricated references, cross-checking of empirical claims, and recognition of algorithmically biased information. These competencies should be assessed through research assignments, seminars, projects, and examinations rather than treated solely as optional digital-skills activities.

3. Faculty-Specific AI-Literacy Interventions

The university should develop faculty-specific AI-literacy interventions that reflect differences in disciplinary practices, students' AI usage patterns, academic levels, and information requirements. Faculties should identify the particular risks associated with AI-generated information within their disciplines and develop tailored training on source evaluation, responsible AI use, citation practices, and detection of hallucinated or biased content. Faculty members should also receive training so that they can guide students effectively in evaluating AI-generated academic materials. Particular attention should be given to groups identified through institutional assessment as having lower AI-literacy levels or greater susceptibility to misinformation. Periodic assessment should be conducted to determine whether these interventions improve students' AI literacy, epistemic trust, and ability to resist misinformation.

4. University Guidelines for Responsible AI Use, Source Verification and Academic Citation

Universities should formulate and regularly update comprehensive guidelines for responsible AI use in teaching, learning, research, and academic writing. The guidelines should clearly specify acceptable and unacceptable uses of generative AI, requirements for disclosure of substantial AI assistance where appropriate, procedures for verifying AI-generated sources and claims, and standards for accurate academic citation. They should also require students and researchers to independently verify references rather than treating AI-generated citations as automatically reliable. The guidelines should incorporate awareness of algorithmic bias and hallucinated citations and provide practical procedures for reporting, correcting, and preventing inaccurate AI-generated academic information. Implementation should be supported by faculty-level monitoring, academic-integrity mechanisms, and periodic review so that institutional policies remain responsive to changes in AI technologies and patterns of student use.

These four recommendations collectively address the study's four objectives by linking AI literacy with epistemic trust and misinformation susceptibility, strengthening awareness of hallucinated citations and algorithmic bias, developing interventions around the identified predictive factors, and accommodating differences associated with gender, AI usage, academic level, and faculty.

REFERENCES

[1] A. Adel and N. Alani, "Can generative AI reliably synthesise literature? Exploring hallucination issues in ChatGPT," *AI & Society*, vol. 40, pp. 6799–6812, 2025, doi: 10.1007/s00146-025-02406-7. Springer

[2] Y. Albadarin, M. Saqr, N. Pope, and M. Tukiainen, "A systematic literature review of empirical research on ChatGPT in education," *Discover Education*, vol. 3, Art. no. 60, 2024, doi: 10.1007/s44217-024-00138-2. Springer

[3] H. Alkaissi and S. I. McFarlane, "Artificial hallucinations in ChatGPT: Implications in scientific writing," *Cureus*, vol. 15, no. 2, Art. no. e35179, 2023, doi: 10.7759/cureus.35179. Crossref

[4] O. Almatrafi, A. Johri, and H. Lee, "A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023)," *Computers and Education Open*, vol. 6, Art. no. 100173, 2024, doi: 10.1016/j.caeo.2024.100173. Crossref

[5] C. Baek, T. Tate, and M. Warschauer, "'ChatGPT seems too good to be true': College students' use and perceptions of generative AI," *Computers and Education: Artificial Intelligence*, vol. 7, Art. no. 100294, 2024, doi: 10.1016/j.caeai.2024.100294. ScienceDirect

[6] M. I. Baig and E. Yadegaridehkordi, "ChatGPT in higher education: A systematic literature review and research challenges," *International Journal of Educational Research*, vol. 127, Art. no. 102411, 2024, doi: 10.1016/j.ijer.2024.102411. ScienceDirect

[7] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?," in *Proc. 2021 ACM Conf. Fairness, Accountability, and Transparency*, 2021, pp. 610–623, doi: 10.1145/3442188.3445922. ACM

[8] B. Billingsley, "Preparing students to engage with science- and technology-related misinformation: The role of epistemic insight," *The Curriculum Journal*, vol. 34, no. 2, pp. 335–351, 2023, doi: 10.1002/curj.190. Crossref

[9] C. Bruce, *The Seven Faces of Information Literacy*. Auslib Press, 1997.

[10] A. Cheng, V. Nagesh, S. Eller, V. Grant, and Y. Lin, "Exploring AI hallucinations of ChatGPT: Reference accuracy and citation relevance of ChatGPT models and training conditions," *Simulation in Healthcare*, vol. 20, no. 6, pp. 413–418, 2025, doi: 10.1097/SIH.0000000000000877. PubMed

[11] K. K. C. Cheung, J. K. H. Pun, and W. Li, "Students' holistic reading of socio-scientific texts on climate change in a ChatGPT scenario," *Research in Science Education*, vol. 54, pp. 957–976, 2024, doi: 10.1007/s11165-024-10177-2. Crossref

[12] V. Hofmann, P. R. Kalluri, D. Jurafsky, and S. King, "AI generates covertly racist decisions about people based on their dialect," *Nature*, vol. 633, pp. 147–154, 2024, doi: 10.1038/s41586-024-07856-5. Nature

[13] K. Y. Hussen, W. T. Sewunetie, A. A. Ayele, S. H. Imam, S. H. Muhammad, and S. M. Yimam, "The state of large language models for African languages: Progress and challenges," *arXiv preprint arXiv:2506.02280*, 2025, doi: 10.48550/arXiv.2506.02280. arXiv

[14] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, "Survey of hallucination in natural language generation," *ACM Computing Surveys*, vol. 55, no. 12, Art. no. 248, 2023, doi: 10.1145/3571730. ACM

[15] M. C. Laupichler, A. Aster, J. Schirch, and T. Raupach, "Artificial intelligence literacy in higher and adult education: A scoping literature review," *Computers and Education: Artificial Intelligence*, vol. 3, Art. no. 100101, 2022, doi: 10.1016/j.caeai.2022.100101. Crossref

[16] J. Lee, Y. Hicke, R. Yu, C. Brooks, and R. F. Kizilcec, "The life cycle of large language models in education: A framework for understanding sources of bias," *British Journal of Educational Technology*, vol. 55, no. 5, pp. 1982–2002, 2024, doi: 10.1111/bjet.13505. Wiley

[17] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human Factors*, vol. 46, no. 1, pp. 50–80, 2004, doi: 10.1518/hfes.46.1.50_30392. Crossref

[18] T. Lintner, "A systematic review of AI literacy scales," *npj Science of Learning*, vol. 9, Art. no. 50, 2024, doi: 10.1038/s41539-024-00264-4. Nature

[19] D. Long and B. Magerko, "What is AI literacy? Competencies and design considerations," in *Proc. 2020 CHI Conf. Human Factors in Computing Systems*, 2020, pp. 1–16, doi: 10.1145/3313831.3376727. ACM

[20] B. D. Lund, Z. A. Teel, Y. Mohammed, A. Jagathpally, and T. Wang, "Artificial intelligence (AI) and information seeking: A comparative exploration of AI chatbots, search engines, and library resources as information sources among university students," *Journal of Librarianship and Information Science*, advance online publication, 2026, doi: 10.1177/09610006261438484. SAGE

[21] B. D. Lund and T. Wang, "Chatting about ChatGPT: How may AI and large language models influence academia and libraries?," *Library Hi Tech News*, vol. 40, no. 5, pp. 26–29, 2023.

[22] M. Z. Naser, "Hallucinations in generative artificial intelligence and large language models: Tests, datasets, detection and correction methods," *Language Resources and Evaluation*, vol. 60, Art. no. 64, 2026, doi: 10.1007/s10579-026-09938-4. Springer

[23] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, "AI literacy: Definition, teaching, evaluation and ethical issues," *Proceedings of the Association for Information Science and Technology*, vol. 58, no. 1, pp. 504–509, 2021, doi: 10.1002/pra2.487. Crossref

[24] J. Ojo, K. Ogueji, P. Stenetorp, and D. I. Adelani, "How good are large language models on African languages?," *arXiv preprint arXiv:2311.07978*, 2023, doi: 10.48550/arXiv.2311.07978. arXiv

[25] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Art. no. n71, 2021, doi: 10.1136/bmj.n71. BMJ

[26] S. Y. Rieh, "Judgment of information quality and cognitive authority on the Web," *Journal of the American Society for Information Science and Technology*, vol. 53, no. 2, pp. 145–161, 2002, doi: 10.1002/asi.10017. Crossref

[27] E. R. Spearing, C. I. Gile, A. L. Fogwill, T. Prike, B. Swire-Thompson, S. Lewandowsky, and U. K. H. Ecker, "Countering AI-generated misinformation with pre-emptive source discreditation and debunking," *Royal Society Open Science*, vol. 12, no. 6, Art. no. 242148, 2025, doi: 10.1098/rsos.242148. PubMed

[28] D. Sperber, F. Clément, C. Heintz, O. Mascaro, H. Mercier, G. Origgi, and D. Wilson, "Epistemic vigilance," *Mind & Language*, vol. 25, no. 4, pp. 359–393, 2010, doi: 10.1111/j.1468-0017.2010.01394.x. Crossref

[29] A. Stojanov, Q. Liu, and J. H. L. Koh, "University students' self-reported reliance on ChatGPT for learning: A latent profile analysis," *Computers and Education: Artificial Intelligence*, vol. 6, Art. no. 100243, 2024, doi: 10.1016/j.caeai.2024.100243. ScienceDirect

[30] N. J. Tanchuk and R. M. Taylor, "Personalized learning with AI tutors: Assessing and advancing epistemic trustworthiness," *Educational Theory*, vol. 75, no. 2, pp. 327–353, 2025, doi: 10.1111/edth.70009. Wiley

[31] W. H. Walters and E. I. Wilder, "Fabrication and errors in the bibliographic citations generated by ChatGPT," *Scientific Reports*, vol. 13, Art. no. 14045, 2023, doi: 10.1038/s41598-023-41032-5. Nature

[32] P. Wilson, *Second-hand Knowledge: An Inquiry into Cognitive Authority*. Greenwood Press, 1983.