Journal of Education, Innovation, and Communication (JEICOM)

DOI: https://doi.org/10.34097/jeicom-8-2-5
Editors: Dr. Margarita Kefalaki & Dr. Fotini Diamantidaki

"We Talk About the Risks, What About the Gains?" Large Language Models and Learning in Mathematics, Statistics, and Computing

Valentine Joseph Owan[1]

ABSTRACT

Large language models are increasingly present in higher education worldwide, yet debates in many institutions still concentrate mainly on fear, integrity concerns, and possible harm to student learning. This paper is a conceptual and argumentative contribution grounded in an extensive review of recent international and African literature on mathematics, statistics, and computational learning. It aims to clarify how large language models can strengthen quantitative learning when supported by responsible policy, ethical guidance, and sound pedagogy. The reviewed literature indicates that these systems can support stepwise reasoning in mathematics, deepen statistical understanding, enhance programming competence, strengthen writing and reporting skills, and improve student motivation, especially where traditional feedback is limited. The paper argues that meaningful benefit depends on AI literacy, prompting competence, guided use, verification practices, and strong lecturer guidance. It further contends that African higher education requires curriculum reform, assessment redesign, and institutional readiness to integrate AI tools responsibly, while also attending to access, equity, and infrastructural realities. The paper concludes that large language models should not only be viewed as threats. When planned carefully, they can support learning quality, widen participation, and strengthen competence in quantitative disciplines that are critical for development in Nigeria and across Africa.

Keywords: Artificial intelligence; quantitative education; mathematics and statistics learning; computational learning; African higher education

[1] University of Calabar, Calabar, Nigeria - owanvalentine@gmail.com

1 BACKGROUND AND PROBLEM FRAMING

Debates about large language models in education have often leaned toward fear, suspicion, and control rather than toward measured academic judgment. Many institutions concentrate on cheating, shortcuts, and the risk that students may lose important skills when using these systems. Scholars repeatedly warn about integrity concerns, uncertain factual reliability, and unstable reasoning in AI-assisted academic work, and these issues remain central to discussions in higher education (Raza et al., 2025). Recent Nigerian studies also report worries about student misuse, including deliberate academic dishonesty and unsafe learning habits (Mohammed et al., 2025; Ofem et al., 2025). However, these cautionary debates sometimes overshadow careful examination of the educational value that these systems already provide, especially in mathematical, statistical, and computational disciplines, where difficulty levels, anxiety, and resource limitations already challenge many learners.

Evidence in the literature indicates that large language models have moved beyond experimental curiosities and now serve meaningful academic purposes. A study reports that they can support reasoning, assist with coding, and provide guided responses to help students manage demanding quantitative tasks (Qu et al., 2025). Nigerian and African evidence further indicates that students increasingly accept AI tools for academic work, especially in self-directed learning environments, when guidance is present and academic integrity expectations remain clear (Owan, Chukwu et al., 2025; Owan, Mohammed et al., 2025). Advances in multimodal systems contribute to this progress because newer tools can process text, visuals, symbolic notation, and other inputs central to mathematics and computing education (Qu et al., 2025; Yin et al., 2024). These abilities create new learning possibilities that traditional lecture-centred environments often struggle to sustain, particularly in large African classrooms where students have limited access to personalised academic support (Ofem et al., 2025).

These developments cannot depend on technological excitement alone. Their value in education requires evidence of learning gains, student growth, and development of confidence and competence (Owan, Chukwu et al., 2025). A focus on educational gains guides this paper, informed by Nigerian and broader African academic realities in which limited resources, unequal access to specialist support, and heavy teaching loads continue to affect quantitative education. The purpose is to bring together reported benefits while acknowledging necessary limits and responsible safeguards (Mohammed et al., 2025; Raza et al., 2025). Evidence from beyond Africa is interpreted in relation to African classrooms, resource conditions, and teaching realities. This paper synthesises relevant scholarly literature and recent studies, underpinned by broader conceptual reasoning, to explain the educational value of large language models while maintaining attention to responsible practice. The literature reviewed in this paper suggests that large language models should not only attract suspicion and punishment. They should be regarded as academic tools whose usefulness depends on careful institutional planning, responsible learner engagement, and sound educational guidance (Owan et al., 2023).

1.1 Approach to literature review and synthesis

This paper adopts a conceptual and narrative review approach grounded in recent literature on large language models in mathematics, statistics, computing, and higher education. The review drew primarily from peer-reviewed journal articles, conference papers, and emerging scholarly reports published between 2023 and 2025, with selected foundational theoretical works included where relevant. Literature was identified through searches of major academic databases and indexing platforms, including Scopus, Web of Science, Google Scholar, ERIC, and publisher databases, using combinations of keywords related to large language models, generative AI, mathematics education, programming education, statistics learning, AI literacy, higher education, and African universities.

Sources were selected based on their relevance to quantitative learning, higher education practice, educational implications of large language models, and discussions connected to pedagogy, reasoning, assessment, or institutional policy. Particular attention was given to studies on mathematics, statistics, computational learning, and African or resource-constrained educational settings. The evidence was synthesised thematically rather than statistically, with emphasis placed on recurring patterns relating to learning gains, instructional support, student engagement, ethical concerns, academic integrity, and institutional readiness.

The review does not claim the comprehensiveness of a formal systematic review or meta-analysis. As a conceptual synthesis, it is limited by potential publication bias, the rapidly evolving nature of AI research, and the uneven availability of African-centred empirical studies. In addition, many cited studies rely on short-term interventions or emerging evidence that may change as technologies and educational practices continue to develop. These limitations were considered when interpreting findings and drawing educational implications.

2 THEORETICAL LENS

Understanding the educational relevance of large language models requires grounding in established theory rather than romantic enthusiasm or fear. Large language models function as cognitive tools that guide learners, reduce effort in demanding tasks, and support engagement with reasoning. Research explains that students benefit when systems provide guided explanations in mathematics, computational support, and incremental tutoring in programming (Kumar et al., 2023). These gains align with cognitive load theory, which posits that learners perform better when unnecessary strain is reduced, and essential thinking is prioritised (Sweller, 1998).

Students in quantitative fields often struggle not because of weak ability, but because symbolic language, complex procedures, and limited feedback place heavy pressure on working memory (Esuong et al., 2022). When large language models clarify steps, present reasoning in stages, or assist with coding breakdown, learners progress more clearly through difficult tasks. Learning improves when this support remains purposeful and supervised (Jošt et al., 2024; Plaat et al., 2025).

A second theoretical base links to constructivist (Dewey, 1938) and inquiry-oriented (Piaget, 1952) learning traditions, which emphasise sense-making, reasoning, and reflection. Recent scholarship indicates that large language models support learning more effectively when educators use them in environments that encourage questioning and careful judgment of AI responses, rather than passive acceptance (Arif & Naeem, 2025). Studies report stronger gains when learners compare AI-generated responses, engage in guided reflection, and discuss reasoning in collaborative settings (Andoko et al., 2025). Evidence from computer science education also points to better outcomes when AI tools stimulate inquiry and analytical dialogue rather than serve as answer sources (Naik et al., 2024). Wider literature explains that learning value depends on educator leadership that promotes scepticism, verification habits, and disciplined engagement, while uncritical reliance may weaken judgement (Yan et al., 2023; Zhui et al., 2024). Nigerian findings support this position, noting that students benefit more when guided to use AI responsibly (Owan, Mohammed et al., 2025). Careful interrogation of AI responses and verification with independent tools encourages deeper learning rather than simple memorisation.

This creates an important theoretical position. Large language models work best when combined with verification tools such as computer algebra systems, statistical software, and coding test environments. Evidence indicates that learning is strengthened when AI provides structured explanation, while independent systems confirm accuracy, supporting logical thinking, improving correctness, and encouraging cautious trust rather than blind reliance (Li et al., 2025; Qu et al., 2025). This approach suits African universities where lecturers cannot always provide detailed individual feedback.

Modern quantitative education also demands reasoning, critique, communication, and practical application rather than narrow recall. Programming requires problem-solving, debugging, collaboration, and development skills. Large language models align with these aims because they generate staged reasoning, simulate problem-solving processes, and restate complex ideas in accessible language, which helps students handle difficult concepts (Plaat et al., 2025; Safranek et al., 2023). When learners first think through problems, attempt solutions, and then use AI to probe explanations and refine their understanding, studies report stronger competence and higher confidence in STEM learning (Matzakos & Moundridou, 2025; Wang et al., 2025). This experience aligns with African teaching realities where guided support helps students remain committed to demanding learning.

3 GAINS IN PROGRAMMING, COMPUTATIONAL THINKING, AND WEB OR SOFTWARE DEVELOPMENT

Programming education remains one of the most demanding areas for university learners, including those in Nigeria and across Africa. Students often struggle with syntax, problem breakdown, debugging frustrations, and limited access to consistent feedback. Many programming laboratories are crowded, lecturers carry heavy teaching loads, and peer support varies in quality. Within this learning situation, large language models now provide meaningful academic support across coding activities, conceptual learning, and practice.

3.1 Performance gains and task support

Research indicates that students who use LLM-based assistants often perform better in programming tasks than those who rely only on traditional resources or unsupported work (Li et al., 2025; Omeh et al., 2025; Park & Kim, 2025). Studies describe gains in task completion, clearer understanding of code structure, and greater problem-solving fluency when learners employ LLMs as supportive companions in their coding processes (Kiesler et al., 2025; Matzakos et al., 2023). Immediate feedback also helps students maintain learning momentum. In many African classrooms where one lecturer manages large cohorts, such personalised support is rarely available through human contact alone.

Beyond finishing tasks, LLMs assist across a wide range of programming needs. Evidence records frequent use for code commenting, debugging, rewriting difficult sections, and translating between programming languages. This reduces common stumbling blocks that often discourage beginners (Kiesler et al., 2025). Many students struggle to locate errors, interpret compiler feedback, or understand reasons for program failure. When an LLM explains faults in clear language or proposes workable corrections, learning becomes less hostile, and students stay engaged for longer.

Empirical work also indicates gains in learning efficiency, as learners spend less time stuck on repetitive technical problems and more time engaging with program logic and concepts (Choi & Kim, 2025; Li et al., 2025). This support matters greatly in Nigerian universities where students already manage time pressure, heavy course loads, and infrastructural challenges. In such environments, tools that reduce wasted effort while improving clarity have clear educational value.

3.2 Patterns of use and learning quality

However, gains do not arise automatically. Research explains that the quality of learning depends strongly on how students engage with LLMs. Students who work critically with system output, modify suggestions thoughtfully, and question recommendations tend to perform better than those who request full solutions without reasoning through the task (Kiesler et al., 2025; Matzakos & Moundridou, 2025). This point matters to policymakers and lecturers who worry that students may replace thinking with ready-made responses.

Learner competence generally strengthens when LLM use supports critical and reflective work rather than merely taking simple answers. Studies in medical and higher education report stronger gains in knowledge and skills when students test AI responses, compare them with trusted sources, and discuss their weaknesses rather than accepting outputs as the final truth (Abd-Alrazaq et al., 2023; Safranek et al., 2023). The central challenge is not the availability of AI tools but the learning culture that guides their use. Clear expectations, ethical guidance, and encouragement of analytical judgement help students to engage responsibly.

Evidence further indicates that students often begin by using LLMs as answer machines, but with guidance, they develop stronger prompting habits, request explanations, check reasoning, and explore alternative approaches. This process builds maturity and confidence (Neshaei et al., 2025; Zhang & Huang, 2024). When educators frame LLMs as partners for reflection, diagnosis, and refinement, learners develop responsibility for reasoning and verification (Kim, 2023). This is particularly important for novice programmers whose early failures can cause anxiety and withdrawal. Guided engagement tends to support more stable and meaningful gains.

This has clear relevance for Nigerian higher education. Many universities receive frequent concerns about shallow learning in programming, where students memorise code fragments without real understanding. Responsible integration of LLMs can strengthen reasoning through interactive explanations, guided debugging, and gradual scaffolding. These practices align with competence-based curriculum reforms in Africa that prioritise observable ability, applied problem-solving, and meaningful performance over narrow examination outcomes (Nzoka, 2024; Wang, 2019).

3.3 Supporting foundational computational thinking

Programming does not stand alone. It sits within the broader framework of computational thinking, including decomposition, abstraction, pattern reasoning, and algorithmic planning. Research indicates that LLMs can support these foundational abilities by guiding learners through explanations, suggesting logical pathways, and allowing experimentation without fear of failure (Kiesler et al., 2025; Mohammed et al., 2025). Seeing alternative approaches improves flexibility of thought, while natural-language explanations help many learners understand programming concepts more clearly.

Many Nigerian students encounter structured computational thinking for the first time at university due to limitations in their earlier schooling. In such situations, a supportive explanation becomes a vital learning bridge. Students also report greater willingness to attempt difficult tasks because assistance is available, which strengthens perseverance. This encourages educational inclusion, especially for learners who may be first-generation university students or come from under-resourced schools.

3.4 Extending gains to web and software development

The benefits of programming learning extend to web and software development. Programming skills now support application development, database systems, web platforms, and software engineering practice. Literature explains that LLMs help students navigate application interfaces, generate initial code structures, provide guidance on front-end frameworks, and support deployment-related activities in academic and semi-professional settings (Jošt et al., 2024). This reduces the psychological barrier between learning in class and building real projects—students who once lacked confidence now submit more comprehensive applications because reliable support is available.

For African learners, this outcome has strategic national importance. Many countries, including Nigeria, seek to build strong local software capacity and expand youth participation in digital economies. LLM support enables students to practise beyond class hours, explore frameworks independently, and apply knowledge to real social and business challenges. Evidence also indicates improved documentation habits, clearer report writing, and better project organisation among students who use LLM guided environments (Neshaei et al., 2025; Zdravkova & Ilijoski, 2025). Such abilities are essential in professional programming environments where clarity and communication are as important as logic.

3.5 Balancing support with educational integrity

While programming gains are clear, responsible practice remains essential. Concerns such as plagiarism, uncritical copying, and the risk of replacing student thinking are valid. However, the literature indicates that when learning activities are carefully designed, benefits outweigh risks. Well-planned assessment, honest dialogue about responsible AI use, and reliable verification measures recur in research as practical ways to protect integrity while still delivering educational value. Scholars discourage blanket bans and advise institutions to redesign assessment, encourage authentic and process-based work such as drafts, oral defences, and projects, and update academic honesty policies to permit AI use within clear boundaries (Perkins, 2023; Tran et al., 2025).

A previous work recommends strong institutional policies and culturally sensitive guidance that recognise AI as both a learning aid and a potential source of misconduct (Yusuf et al., 2024). Another strand emphasises AI literacy, ethical training, and structured discussions on responsible practice, including expectations for disclosure, citation, and thoughtful use rather than silent dependence (Malik et al., 2024; Pratama, 2025). Reviews also call for verification systems and academic supervision, in which AI outputs are checked against reliable sources and monitoring tools are applied judiciously to protect fairness and trust (Chansa et al., 2025; Tyndall et al., 2025).

Across this literature, academic integrity depends on sound assessment design, ethical guidance, and disciplined verification rather than fear-driven restriction. Programming education therefore requires balanced integration strategies that protect integrity while supporting learning gains. Instead, it should move forward, encourage transparency in AI-supported learning, and strengthen expectations for reasoning, originality, and explanation. This balanced path supports competence while maintaining ethical rigour.

4 GAINS IN MATHEMATICAL REASONING AND MODELLING

Mathematics education in many Nigerian and African higher institutions remains demanding due to cumulative learning gaps, limited access to tutoring, and a heavy reliance on lecture-centred instruction. Learners frequently struggle with abstract reasoning, algebraic manipulation, and multi-step problem solving. These challenges discourage many students from progressing into statistically and computationally intensive disciplines. In this reality, large language models offer an opportunity to support mathematical reasoning, deepen conceptual clarity, and strengthen modelling capabilities when used carefully and responsibly.

4.1 Step-by-step explanations and reasoning support

A major educational strength of large language models in mathematics lies in their ability to provide guided explanations rather than just final answers. Evidence indicates that LLM-supported tutoring strengthens learning when it prioritises sequential reasoning, clear strategy explanation, and personalised guidance. Studies report gains in comprehension, engagement, and retention when learners receive reasoning chains and tailored explanations rather than bare solutions (Kumar et al., 2023; Ye et al., 2024). Systems that verify thinking at each stage also help students notice where errors arise, which leads to more accurate feedback and a stronger grasp of problem-solving logic. The broader mathematics education literature links segmented reasoning and corrective feedback to improved performance, greater confidence, and reduced error anxiety (Shang et al., 2023; Shelton, 2016). These findings suggest that exposure to clear explanations, instead of quick answers, helps students internalise strategies and transfer knowledge to new tasks.

This benefit carries particular significance for African and Nigerian learning environments where mathematics teaching has often relied on procedural memory, strict attention to final answers, and limited tolerance for error. Many Nigerian learners grow up with mathematics associated with fear and harsh judgment. Stepwise LLM explanations present mathematics as a thinking journey rather than a space for punishment. When reasoning paths become visible, students understand strategies better, recognise relationships across concepts, and gain confidence to persist.

This form of support also fits African classrooms where lecturers often manage large classes with limited resources and little time for one-to-one explanation. Many students leave lessons with unresolved confusion, not because they lack ability, but because they do not receive detailed breakdowns of difficult steps. LLM explanations add another layer of academic support that complements teachers' efforts. Lecturers still guide learning and provide judgment, while LLMs assist with repeated explanation, hints, and alternative reasoning pathways at scale. In this way, careful use of LLMs supports wider access, greater confidence, and deeper mathematical understanding in Nigerian and African higher education.

4.2 Mathematical rule learning and generalisation

Beyond explanation, large language models now display growing ability in mathematical reasoning through learning, applying, and generalising mathematical rules. Research indicates that these systems can internalise principles such as distributive reasoning, equation manipulation, and algebraic generalisation, and reuse them in new problem situations (Lake & Baroni, 2023; Webb et al., 2022). This provides learners with varied examples and multiple reasoning routes, which is especially helpful for students who need repeated exposure before mastery develops.

Evidence further suggests that mathematical reasoning performance improves as models undergo reinforcement learning and advanced training processes. Studies report that newer systems tackle more complex problems more effectively when they simulate reasoning rather than guess final answers (Plaat et al., 2025; Zhong et al., 2026). This points to steady advancement rather than fixed capability. For African education, this matters. AI competence in mathematical reasoning is not static, and current systems already provide meaningful help that will likely strengthen as development continues.

This steady improvement is important for African higher education, where many institutions still face overcrowded classrooms, limited tutorial support, and scarce mathematics laboratories. In such environments, LLMs can function as supplementary reasoning companions, providing consistent practice, reinforcement, and explanation. This support does not replace lecturers. Instead, it creates more learning opportunities for students who would otherwise have limited guided practice.

4.3 Supporting modelling, problem translation, and analytical thinking

Mathematics in higher education extends beyond symbolic manipulation. It prepares students to interpret real situations, model relationships, and analyse problems carefully. Recent literature explains that large language models support this by helping learners move from natural-language descriptions to formal mathematical representations. They guide students in identifying relevant quantities, clarifying assumptions, and framing problems in a workable form (Matzakos & Moundridou, 2025; Zhou, 2025). This matters because many learners find it easier to solve equations than to construct them. Repeated exposure to LLM-guided reasoning strengthens modelling awareness and the ability to formulate new problems independently.

In learning statistical and applied mathematics, LLMs also help students understand the meaning behind the procedures. They clarify what results imply, explain the logic behind models, and relate outcomes to real situations. Nigerian universities increasingly emphasise applied mathematics for engineering, agriculture, economics, and social sciences. Tools that deepen reasoning and interpretation, therefore, align well with national and continental development needs.

Large language models also broaden problem-solving flexibility by presenting alternative solution strategies rather than a single fixed route. When more than one method exists, they explain each approach, encourage comparison, and guide learners to judge strengths and weaknesses. This broadens analytical thinking and reduces narrow dependence on routine procedures (Matzakos & Moundridou, 2025; Zhou, 2025). Such flexibility is vital for research, industrial applications, and policy analysis.

4.4 Benefits for confidence, persistence, and learning attitude

Mathematics anxiety remains a recognised barrier to learning in Nigeria and other regions. It contributes to avoiding mathematically demanding courses and reduces confidence in problem-solving. Many students carry an early fear of mathematics into university, where hesitation and self-doubt continue. Large language models can support emotional readiness for learning by providing immediate explanations and repeated practice without public embarrassment. Studies on AI-supported mathematics learning report gains in self-belief, positive emotional response, and engagement when students use AI-based explanations with regular instruction (Lademann et al., 2025; Ye et al., 2024). Learners can ask questions freely, revisit explanations at their own pace, and check understanding privately. As anxiety reduces, participation and persistence tend to improve, and students report greater confidence (Canonigo, 2024; Lademann et al., 2025).

This emotional benefit directly connects to widening participation aims in African higher education. When learning environments reduce intimidation and provide accessible guidance, more students can participate meaningfully in mathematically intensive programmes. Evidence from AI-supported mathematics settings indicates that learners from less advantaged backgrounds can strengthen understanding and self-belief when they receive supportive explanations rather than only high-stakes grading (Lademann et al., 2025; Ye et al., 2024). In such situations, students who might withdraw from demanding courses are more likely to remain and succeed.

4.5 Recognised Limitations and Responsible Interpretation

While benefits are meaningful, responsible interpretation requires acknowledgement of limits. Studies indicate that large language models sometimes produce incorrect reasoning, numerical mistakes, or weak performance on highly abstract or unfamiliar problems (Fang et al., 2025; Owan et al., 2023). Blind trust, therefore, remains unsafe in high-stakes academic work. Educational use must include verification, critical checking, and human judgment. Blending LLMs with established mathematical verification systems remains necessary, and learners need guidance on assessing responses and validating results. When such a balance exists, limitations become manageable learning challenges rather than grounds to reject technological assistance entirely (Canonigo, 2024).

In summary, research suggests that large language models can strengthen mathematical reasoning, deepen conceptual understanding, support modelling ability, and encourage perseverance. Their value increases when educators combine them with verification strategies and clear guidance. In Nigerian and African higher education, this capacity presents a meaningful opportunity to expand access, enhance competence, and build confidence among learners who often struggle with mathematics.

5 GAINS IN STATISTICAL LITERACY, DATA ANALYSIS, AND STATISTICAL WRITING

While mathematics education often emphasises symbolic reasoning and modelling, statistics education places greater emphasis on interpretation, uncertainty, inference, and communication. These disciplinary differences influence how large language models support learning within statistical education. Statistics forms a central pillar of modern education in science, social science, health, technology, and development research. However, across many African institutions, including Nigeria, students often approach statistics with anxiety and uncertainty. Many struggle with probability, hypothesis testing, regression interpretation, and formal reporting of results. Limited tutor availability and crowded classrooms frequently intensify these challenges. Within such conditions, large language models provide meaningful support for understanding, application, and communication of statistical knowledge.

5.1 Strengthening Conceptual Statistical Understanding

Statistical literacy requires a clear understanding of what key measures represent and how they inform decisions. Evidence indicates that AI systems can improve comprehension of p-values, confidence intervals, correlation, and regression modelling by providing natural language explanations and supporting interactive questioning where students receive immediate responses (Song et al., 2025). This guidance is important because many learners have traditionally relied on memorisation without sufficient interpretation. They may know how to calculate values but remain unsure how to reason with the results. Through an LLM-supported explanation, statistical outcomes become clearer. Learners begin to understand what a confidence interval suggests about uncertainty, what statistical significance represents as evidence, and how regression coefficients relate to real-life relationships. These improvements strengthen confidence and analytical clarity.

Interactive systems also allow repeated questioning in a safe learning environment. Students can revisit difficult ideas at a suitable pace without fear of embarrassment or impatience. This form of support is especially valuable in African universities where lecturer consultation time is limited, and peer guidance is unreliable. In addition, LLMs help learners connect theory with application. When students conduct statistical tests, the system can explain why a specific test is appropriate, which assumptions apply, and what alternative approaches may entail. In this way, learning moves beyond mechanical computation towards meaningful reasoning (Matzakos & Moundridou, 2025). In applied fields such as public health, agriculture, economics, and the social sciences, this kind of understanding is essential for responsible professional practice.

5.2 Support for data analysis practice and computational engagement

Statistics is not only concerned with theory. It also requires practical engagement with data. Learners must analyse datasets, build models, interpret findings, and explain results in clear language. Recent work indicates that AI-supported learning environments help students handle datasets, use simulation tools, and engage more confidently with statistical programming. This encourages practice and reduces the sense of being overwhelmed (Matzakos & Moundridou, 2025; Song et al., 2025). Students benefit when AI systems guide them through stages of analysis, recommend visualisations, suggest strategies, or warn about overlooked issues. Instead of struggling alone or abandoning tasks, learners proceed with clearer direction. This support is important in Nigeria, where many universities now prioritise applied data analysis competence to support national development in public health, economics, education planning, and environmental management.

AI tools also encourage exploratory learning. Students can adjust parameters, compare methods, and review alternative outputs without fear of failure. This flexibility promotes curiosity, strengthens research orientation, and expands exposure to analytical techniques beyond limited classroom time. Such an opportunity is particularly valuable for students who lack access to advanced software training due to financial or infrastructural constraints. However, these gains remain strongest when AI support is combined with human teaching and firm academic discipline. Large language models function best as complements rather than replacements. When learners combine AI explanation with recognised statistical software, they benefit from an integrated learning process in which one tool deepens reasoning while the other supports technical accuracy.

5.3 Enhancing statistical writing, communication, and reporting

A major challenge for many postgraduate students lies not in performing statistical analysis, but in communicating results in clear academic language. Learners often produce technically sound tables yet struggle to transform numerical findings into meaningful interpretations. Recent research indicates that large language models provide important support in this area. They help students explain findings more clearly, refine formal writing, and organise sections such as methods, results, and discussion with stronger clarity and purpose (Zdravkova & Ilijoski, 2025). Evidence suggests that such support builds confidence among students with strong analytical abilities but who remain unsure about academic expression.

This benefit is particularly valuable in African universities, including Nigeria, where many capable students work in English as an additional language. Studies indicate that AI-assisted writing tools help learners express complex ideas more coherently, improve sentence accuracy, and reduce ambiguity in interpretation (Granjeiro et al., 2025; Meyer et al., 2023). Large language models also strengthen paragraph cohesion, improve logical flow, and support clearer explanations of analytical reasoning, including justifications of statistical choices and interpretations of outputs (Owan et al., 2023). In research-driven environments, this reduces anxiety around writing and encourages higher-quality scientific communication.

Beyond academic work, AI-supported writing strengthens professional readiness. Graduates across Africa increasingly need to communicate data-based arguments to technical experts, policymakers, and non-specialist audiences. Studies indicate that AI-assisted writing environments improve readability, grammatical precision, and clarity, especially among multilingual learners (Granjeiro et al., 2025; Meyer et al., 2023). In this sense, LLM-enabled writing support does not replace intellectual effort. Instead, it strengthens students' ability to communicate knowledge responsibly and confidently, supporting meaningful participation in research, policy development, and national development.

Although programming, mathematics, and statistics possess distinct learning demands, the literature reviewed across these domains reveals several shared educational patterns. In all three areas, large language models appear most beneficial when they support staged explanation, guided explanation, reflective engagement, and independent checking rather than passive answer retrieval. At the same time, the specific form of support differs across domains. Programming education benefits strongly from debugging assistance and computational scaffolding, mathematics learning gains more from structured reasoning and modelling support, while statistics education benefits particularly from interpretative explanation, analytical guidance, and support for statistical communication. Distinguishing these domain-specific contributions is important because effective educational integration depends on the learning goals, disciplinary practices, and assessment demands within each field.

Table 1: Educational gains, conditions for effective use, and key risks of large language models across quantitative disciplines

Domain	Major Educational Gains	Pedagogical Contributions	Conditions for Effective Use	Key Risks and Limitations
Programming and Computing	Improved debugging, code explanation, task completion, computational thinking, software development support, coding efficiency	Supports scaffolded learning, interactive problem-solving, algorithmic reasoning, project development, and self-directed experimentation	AI literacy, guided prompting, reflective coding practice, verification using compilers/testing tools, lecturer supervision	Overreliance on generated code, plagiarism, shallow conceptual understanding, reduced independent debugging ability, inaccurate code suggestions
Mathematics and Modelling	Stepwise reasoning support, conceptual clarification, modelling assistance, mathematical rule generalisation, improved confidence and persistence	Encourages structured reasoning, multiple solution pathways, problem translation, analytical flexibility, and reduced mathematics anxiety	Verification using symbolic or computational tools, guided interpretation, human oversight, reflective engagement with solutions	Incorrect reasoning chains, hallucinated solutions, weak performance on abstract problems, blind trust in outputs
Statistics and Data Analysis	Improved statistical interpretation, support for data analysis, statistical writing assistance, analytical explanation, enhanced research communication	Strengthens interpretation of statistical concepts, exploratory learning, analytical reflection, and communication of findings	Integration with recognised statistical software, critical evaluation of outputs, lecturer guidance, cross-validation	Misinterpretation of statistical results, inaccurate analytical suggestions, superficial understanding of assumptions, dependence on automated interpretation
Shared Cross-Domain Benefits	Increased engagement, improved motivation,	Encourages active learning, reflection, questioning, and	Responsible AI policies, prompting competence,	Academic dishonesty concerns, unequal access,

	personalised support, faster feedback, and greater academic confidence	independent practice	ethical guidance, digital access, and institutional readiness	infrastructural barriers, and reduced critical thinking when poorly guided

Table 1 indicates that although all three domains benefit from personalised explanation, faster feedback, and reflective engagement, the dominant educational contribution differs across disciplines. Programming gains are closely linked to debugging and computational scaffolding, mathematics gains centre more on structured reasoning and modelling, while statistics benefits strongly from interpretative explanation and communication support. The synthesis also suggests that educational value depends consistently on confirmation of accuracy, structured engagement and institutional support across all domains. Beyond discipline-specific gains, the literature also associates AI-supported learning with broader motivational and engagement-related outcomes across quantitative education.

6 MOTIVATION, ENGAGEMENT, AND ACADEMIC EFFICIENCY

Learning thrives when students remain motivated, confident, and willing to continue despite difficulty. Quantitative subjects often intimidate learners due to perceived difficulty, heavy workloads, and fear of failure. Many lose interest early, especially when feedback is scarce or frustration becomes constant. Under such conditions, large language models can positively contribute to motivation, engagement, and a sense of productivity.

6.1 Improved Engagement and Satisfaction

Meta-analytic and survey findings associate AI-supported learning with higher motivation, particularly in computing and STEM fields (García-Martínez et al., 2023; Mohamed et al., 2025). Students frequently report that AI assistance improves their ability to complete demanding tasks, understand complex problems, and engage more confidently with challenging work. These perceived gains often relate to greater satisfaction with learning experiences, as students feel more supported, more capable, and less overwhelmed by academic difficulty (Dogaru et al., 2025; Mohamed et al., 2025).

This pattern is especially important in African higher education, where many learners manage academic stress alongside financial pressure, infrastructural barriers, and limited access to personalised support. When AI tools reduce unnecessary frustration and create regular experiences of success, students tend to feel more competent and more willing to persevere through demanding STEM pathways (García-Martínez et al., 2023). In this way, increased confidence supports persistence and reduces the likelihood of withdrawal due to discouragement or pressure.

Students also tend to report greater satisfaction as they gain competence in engaging with AI tools. Studies on AI-supported learning environments indicate that improved prompting skills and strategic tool use are often associated with stronger perceptions of usefulness and more positive learning attitudes. Over time, many learners move from simple experimentation to deliberate, purposeful use as AI support becomes part of regular study practice rather than a temporary novelty (Dogaru et al., 2025). This gradual movement towards skilled and responsible use supports more stable motivation, deeper engagement, and stronger academic resilience.

6.2 Productivity, time efficiency, and academic progress

Large language models also contribute to a sense of academic efficiency. They help students complete preparatory stages more quickly, resolve confusion earlier, and make steady progress on assignments. Learners often report feeling more productive, better able to manage complex tasks, and more willing to undertake extended project work when AI support is available (Kiesler et al., 2025; Matzakos et al., 2023). In environments where deadlines are strict and academic demands are heavy, productivity-enhancing tools can provide relief and encourage deeper engagement rather than resignation. Students can then devote more time to understanding principles, refining reasoning, and improving creativity.

This development aligns well with the needs of Nigerian higher education, where building scientific and technological competence remains an essential national goal. When learners remain engaged, progress confidently, and manage academic demands more effectively, higher education systems are better positioned to fulfil their educational and developmental responsibilities. Despite these reported gains, educational value does not emerge automatically from AI access alone.

7 CONDITIONS FOR PRODUCTIVE USE: AI LITERACY, PROMPTING COMPETENCE, AND HUMAN OVERSIGHT

The educational gains associated with large language models in quantitative disciplines do not appear automatically. Evidence indicates that the value of these systems depends strongly on how students use them, how universities organise engagement, and how lecturers guide learning. Large language models reward thoughtful use but can encourage shallow learning when used carelessly. For this reason, competence, instructional planning, and oversight remain central to responsible integration of AI tools in education.

7.1 AI Literacy as Foundational Academic Competence

Studies explain that educational benefits from large language models arise when learners understand how these systems function, the strengths they provide, and the limits that remain. AI literacy underpins responsible use because students need to recognise that large language models can generate convincing responses that may still contain inaccuracies. Gains tend to increase when learners question outputs, verify claims with independent sources, and interpret responses critically rather than accepting every response as authoritative truth (Kiesler et al., 2025; Matzakos & Moundridou, 2025). In this sense, AI literacy now stands beside information literacy, research ethics, and digital proficiency as an essential part of modern academic preparation.

AI literacy involves understanding how AI operates, how to apply it effectively in learning, and how to judge its risks, weaknesses, and ethical concerns (Salhab, 2024; Southworth et al., 2023). Such competence supports balanced and responsible use because informed learners can question responses, recognise bias, and avoid excessive dependence on the system (Rojo et al., 2025; Southworth et al., 2023). Literature on media and information literacy explains that students require discernment, ethical awareness, and critical judgment to benefit from AI tools without losing independence or accepting misinformation uncritically (Rojo et al., 2025).

This requirement carries particular weight in African higher education, where many students encounter advanced AI tools for the first time at the university level. Without guidance, learners may either over-trust AI responses or reject AI completely. University systems that embed AI literacy and ethics into policy and curriculum create better conditions for the balanced use of AI. Research increasingly calls for institution-wide AI literacy strategies that treat AI competence as a core twenty-first-century academic skill, closely linked with data capability, critical thinking, and ethical engagement (Salhab, 2024; Southworth et al., 2023). Recent work also stresses the need for explicit AI governance, including clear policies on academic integrity, ethical use, privacy, and accountability, supported by training for both students and staff (Aler Tubella et al., 2023; Song, 2024). Where such frameworks exist, AI tools are more likely to strengthen learning rather than create confusion or misconduct (Song, 2024).

7.2 Prompting skill and interaction quality

Educational value from large language models depends not only on access to technology but also on how students communicate with it. Evidence from higher education indicates that learners who receive training in effective prompt design gain clearer explanations, more relevant feedback, and stronger learning outcomes than users who receive no guidance (Ramadana et al., 2025; Yang et al., 2025). Weak prompting often leads to vague or oversimplified responses, while deliberate prompting techniques encourage clearer explanations, stepwise guidance, and deeper reasoning (Picardal, 2025; Shen et al., 2025).

Prompting competence closely resembles questioning competence in traditional classroom learning. Research on smart prompting explains that students who learn to ask precise questions, specify constraints, request justification, and seek clarification engage more actively in analysis and evaluation and display richer cognitive effort (Ramadana et al., 2025; Shen et al., 2025). Instruction on effective interaction with AI tools, therefore, strengthens thinking as well as tool output. It promotes clarity, precision, reflection, and disciplined reasoning, which remain central to serious academic work.

Emerging evidence further links improvements in prompting skills to stronger academic performance and higher satisfaction in AI-supported learning environments. More sophisticated prompts and thoughtful follow-up questions are positively associated with course performance and sustained engagement (Ramadana et al., 2025; Yang et al., 2025). Students who develop prompting competence tend to treat AI systems as thinking resources for planning, explanation, and revision rather than shortcuts for ready answers (Picardal, 2025).

For African institutions, integrating prompting education represents a low-cost but high-impact intervention. Universities can introduce short training modules, guided practice tasks, and policy guidance that encourage deliberate questioning, reflective review of AI responses, and responsible engagement with generative tools. Such measures strengthen educational culture, promote ethical use of AI, and help students employ large language models as genuine intellectual partners rather than as convenience tools.

7.3 Guided use and human oversight

Research explains that supervised integration of large language models supports stronger learning outcomes than unguided or fully independent interaction. Students gain more when lecturers design activities that require initial personal effort, followed by structured AI consultation, and then reflective revision. This pattern encourages deeper thinking because learners first attempt solutions, then test ideas with AI support, and finally refine understanding through clarification and feedback (Kiesler et al., 2025; Matzakos & Moundridou, 2025). In such conditions, AI assists thinking instead of replacing it.

Human oversight remains essential. Lecturers, tutors, and institutions need to set clear expectations for responsible AI use, create time for reflection, and ensure that students do not bypass reasoning entirely. AI functions best as a teaching aid that supports explanation, feedback, and practice, while learners remain responsible for understanding and judgment. In this way, AI strengthens learning without removing the importance of human teaching.

This approach is particularly relevant in Nigerian and wider African higher education, where lecturers often manage large student numbers and cannot always provide individual guidance. Guided AI integration can extend instructional reach while preserving academic authority and expertise. By positioning large language models as partners in teaching, universities can expand access to explanation, improve learning support, and still maintain strong academic standards.

7.4 Verification and complementary tool use

Large language models provide valuable explanation support, but they are not flawless. Scholars have identified occasional inaccuracies, reasoning gaps, and unstable performance on highly technical or unfamiliar mathematical problems (Fang et al., 2025). For educational reliability and depth of learning, students need training to verify AI-generated responses using independent tools and established academic procedures. Blending large language models with verification systems, such as computer algebra platforms, statistical software, and programming test frameworks, enhances accuracy and learning safety. One tool supports reasoning, while the other secures correctness.

This partnership encourages disciplined academic behaviour. Students learn to evaluate evidence, cross-check claims, and maintain caution when working with automated systems. This practice aligns with long-standing academic traditions across Africa that value thoroughness, careful reasoning, and respect for evidence. Encouraging verification also prepares students for professional environments where accuracy, accountability, and reliability carry serious consequences.

The literature reviewed in this paper suggests that educational gains depend more on pedagogical and institutional conditions than on technology alone. Figure 1 presents a conceptual synthesis of these relationships by illustrating how AI literacy, prompting competence, supported implementation, cross-verification, and instructional monitoring interact to support responsible educational outcomes in quantitative disciplines.

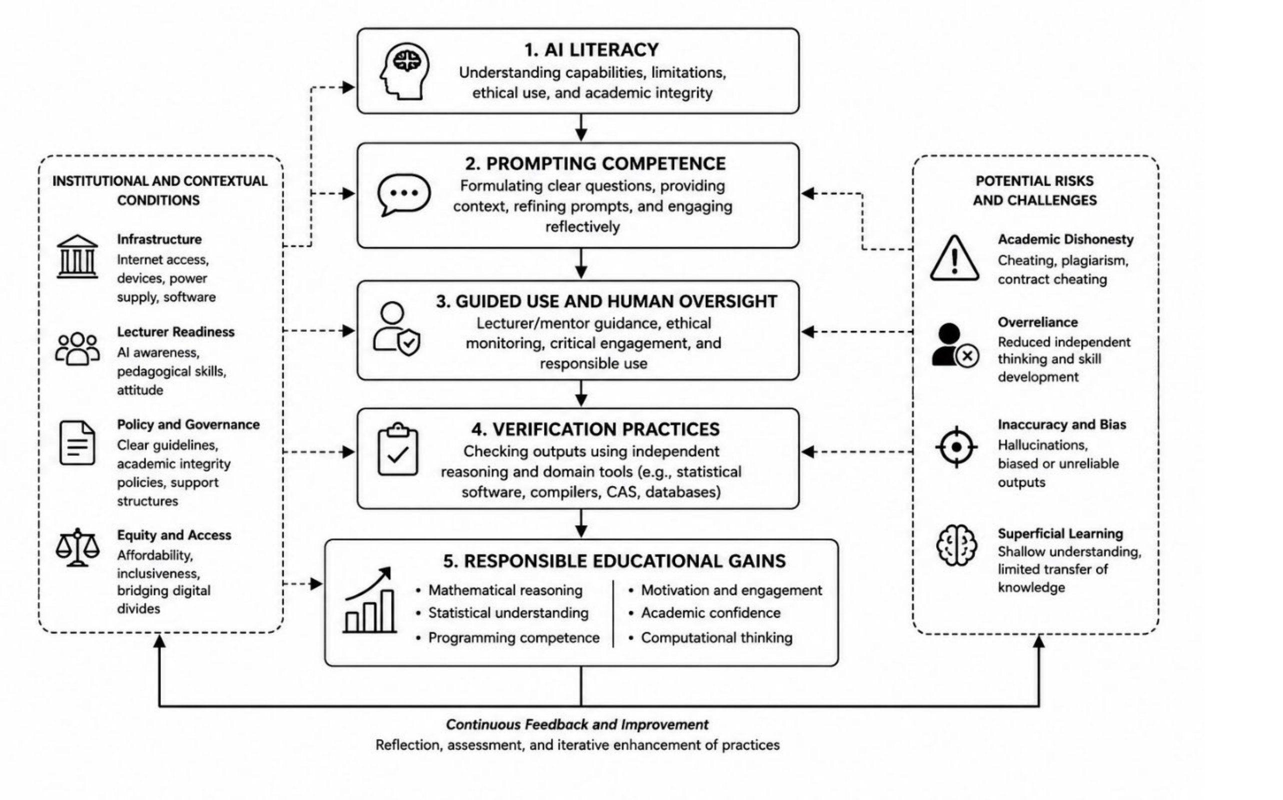

Figure 1. Conceptual relationship among AI literacy, prompting competence, guided use, verification practices, and responsible educational gains in mathematics, statistics, and computing.

Figure 1 emphasises that meaningful educational benefit arises from structured, reflective engagement with AI systems rather than from unrestricted automation. The framework further suggests that institutional readiness, lecturer guidance, and validation practices function as stabilising conditions that help learners gain educational value while reducing risks associated with overreliance, inaccuracy, and superficial learning. Although the literature identifies substantial educational promise, meaningful integration also requires careful management of academic and ethical risks.

8 MANAGING RISKS WITHOUT LOSING EDUCATIONAL GAINS

While large language models carry significant educational promise, responsible academic practice must recognise associated risks. Concerns about academic dishonesty, reduced independent thinking, shallow engagement, bias, and factual instability remain valid and require careful attention. However, evidence suggests that these concerns can be managed effectively through thoughtful educational design rather than reactionary prohibition.

8.1 Academic integrity and responsible use

Academic integrity concerns remain among the most frequently discussed issues. Some fear that large language models enable students to submit work they did not produce meaningfully. Evidence indicates that responsible policy frameworks, clear expectations, and honest assessment strategies can reduce misuse while still allowing educational benefits (Kumar et al., 2023; Raza et al., 2025). Instead of banning AI, institutions can require explanation of processes, justification of reasoning, and transparent documentation of how AI contributed to academic work. Oral assessments, reflective commentary, supervised tasks, and process-based marking encourage genuine learning and discourage shallow copying. Integrity-strengthening assessment practices already exist within African educational systems and can adapt naturally to AI-supported environments.

In addition, AI detection and anomaly analysis tools, although imperfect, can help identify suspicious patterns. However, such tools should support rather than replace sound pedagogy and trust-building academic culture (Matzakos & Moundridou, 2025). The central emphasis should remain on transparency, guidance, and shared responsibility.

8.2 Preventing superficial learning and over-reliance

A further concern is that students may rely excessively on AI tools and fail to develop deep knowledge. The literature suggests that this risk becomes serious primarily in learning environments without guidance, reflection expectations, or verification requirements (Kiesler et al., 2025; Matzakos & Moundridou, 2025). In contrast, when institutions embed structured use, encourage independent thinking before consultation, and require reflection, learning outcomes tend to improve rather than decline. Educational leaders in Africa should therefore resist narrow narratives that present AI only as a danger. Instead, they need to build responsible learning cultures where AI supports understanding rather than replacing thinking. Clear policies, honest classroom dialogue, and lecturer modelling of responsible engagement help learners develop disciplined academic behaviours.

8.3 Reliability, bias, and accuracy concerns

Large language models can sometimes produce inaccurate or biased responses. These risks require institutional awareness and teaching strategies that prepare students to manage them. When learners understand that AI can occasionally generate errors or unfair reasoning, they approach its use with informed caution rather than blind dependence. Evidence indicates that combining AI support with independent verification tools and structured oversight reduces the impact of such weaknesses on learning (D'Souza et al., 2024; Zhang et al., 2025). This disciplined approach prepares students for professional environments in science, engineering, public policy, economics, and health, where automated systems increasingly influence decision-making. Learning to evaluate AI critically, therefore, functions not only as an educational safeguard but also as an essential competence for future careers. The educational implications discussed throughout this paper also carry important consequences for curriculum, assessment, institutional policy, and educational equity.

9 WIDER IMPLICATIONS AND POLICY CONSIDERATIONS

The evidence across programming, mathematics, statistics, and wider computational learning indicates that large language models hold meaningful educational promise when integrated carefully. They enhance learning effectiveness, confidence in reasoning, and engagement. However, these benefits depend on supportive educational policy, institutional readiness, and curriculum redesign. Curricula in mathematics, statistics, and computing now require deliberate recognition of AI as part of the learning environment rather than as an external threat. Research explains that students gain more when AI use forms part of structured learning rather than existing in hidden or forbidden spaces. Curriculum designers, therefore, need to include AI literacy, critical reasoning with AI tools, and responsible prompting competence within formal learning expectations. Assignments should encourage explanation, reflection, and independent thinking rather than focusing only on correct final answers. Empirical work indicates that tasks requiring the justification of decisions, the critical commentary on AI outputs, or the comparison between AI-generated and student reasoning support deeper understanding and stronger metacognitive skills (Liu et al., 2021; Neshaei et al., 2025).

Teaching practice must move in the same direction. Findings indicate that learners behave more responsibly and critically when required to interrogate AI responses, identify weaknesses, and revise reasoning rather than accept machine output unquestioningly (Essien et al., 2024; Zhai et al., 2024). These designs help students maintain intellectual independence while still benefiting from AI assistance. This direction aligns with competence-based reforms across African higher education, which prioritise authentic learning tasks, problem-solving, and evidence of learning processes over reliance on final scores alone (Kikasu et al., 2025). In such systems, lecturers remain central and need professional development to understand AI capability, pedagogy, and ethical use, because institutional leadership and strong teaching cultures remain essential for responsible academic practice.

Assessment processes also require renewal to respond to AI-supported learning. Traditional closed-book procedural examinations alone cannot sustain educational credibility in environments where AI support is widely available. Evidence suggests that assessment should value reasoning, explanation, creativity, and process documentation, not only final correctness (Kiesler et al., 2025; Matzakos & Moundridou, 2025). Practical options include monitored in-class problem solving, oral defences, reflective commentary on AI-assisted tasks, submission of drafts and reasoning trails, and greater emphasis on authentic project work. These approaches do not reduce academic standards. Instead, they strengthen responsibility, originality, and conceptual understanding. Such approaches also align with long-standing African higher education traditions that value honesty, credibility, and disciplined scholarship. Institutions must also develop a clear policy direction on acceptable uses of AI. Ambiguity encourages misuse, while clarity supports responsible learning. Clear communication on legitimate AI assistance, expectations for acknowledgement, and defined boundaries will help protect integrity while supporting meaningful educational benefit.

Large language models can also support equity in African higher education. Many learners across the continent still lack access to high quality tutoring, especially in rural areas, crowded public universities, and underfunded institutions. AI systems can reduce these inequalities by providing immediate explanation, repeated guidance, and supportive practice at relatively low cost (Matzakos & Moundridou, 2025; Matzakos et al., 2023). However, these benefits depend on reliable digital infrastructure, internet access, and affordable devices. Without such investment, AI may widen inequalities instead of reducing them. Policymakers therefore need to treat AI integration as part of national investment in human capital, through stronger digital policy, reliable bandwidth, and institutional support. Cultural considerations also matter. African education traditionally values communal support, mentorship, and shared responsibility for learning. When implemented wisely, AI can extend this supportive ethos by adding another layer of academic assistance. It should strengthen learning communities rather than replace them. Although international evidence strongly informs the current understanding of the use of large language models in education, African-centred empirical studies remain comparatively limited. Consequently, some of the educational implications discussed in this paper represent a cautious interpretation and contextual adaptation of broader international findings rather than conclusions derived exclusively from African evidence.

10 RESEARCH GAPS AND FUTURE DIRECTIONS

Despite the promising findings discussed throughout this paper, interpretation of the evidence requires caution. Much of the existing literature is drawn from short-term interventions, experimental settings, or institutions with stronger digital infrastructure than many African universities. Educational outcomes may therefore vary across disciplines, institutional cultures, resource conditions, and levels of student preparedness. In addition, several reported gains rely heavily on guided implementation, lecturer support, and access to complementary verification tools. These conditions are not yet uniformly available across higher education systems in Africa. For this reason, claims about effectiveness should be interpreted within specific educational and contextual realities rather than assumed to apply universally.

Although strong evidence supports many educational gains, several important gaps in the literature must be acknowledged to inform future research. Much of the current knowledge comes from short evaluation cycles, non-African research environments, or small-scale classroom studies. African-centred evidence is still limited, so there is a need for research grounded in regional realities and long-term outcomes. A key area that requires deeper inquiry concerns long-term learning and knowledge retention. Many studies on AI-supported learning focus on short-term outcomes, such as immediate test improvements or task performance, often without examining whether these gains persist when AI assistance is removed (Bauer et al., 2025; Feigerlova et al., 2025). Longitudinal research is required to track how AI-supported instruction influences sustained retention, transfer of knowledge to new situations, and the durability of conceptual understanding over extended periods. Existing reviews and experimental work stress the importance of determining whether early improvements translate into stable competence and meaningful application in real learning environments (Chance, 2025; Chen, 2025; Herbawani et al., 2025).

Another important direction concerns domain-specific and specialised AI systems. Many educational applications still rely on general-purpose language models that were not originally developed for learning environments. Recent work indicates that domain-tuned tools, for example, systems adapted for mathematics, statistics, or programming instruction, can more accurately capture disciplinary structures and often provide safer, more reliable reasoning (Luo & Yang, 2024; Yousef et al., 2024). Specialised systems created through domain-adaptive training or fine-tuning may therefore provide more stable educational value than unmodified general models (Shi et al., 2025).

A further, highly significant priority is African-centred empirical and policy research. Education systems across Africa face particular realities, including infrastructural constraints, uneven institutional readiness, and persistent digital divides that influence how artificial intelligence can function in real classrooms and universities (Muringa, 2025; Saal et al., 2025). In many countries, unreliable internet connectivity, limited access to devices, and rural-urban inequality restrict meaningful AI adoption. Teaching capacity also varies, as many educators lack AI literacy and structured professional development, which contributes to hesitant or ineffective engagement. Sociocultural conditions, linguistic diversity, and community expectations further influence how students and lecturers respond to AI tools in countries such as Nigeria and South Africa (Essien et al., 2024; Olayinka et al., 2024).

Across the literature, there is growing agreement that African-led research is essential to understand how AI functions within these realities. Grounded evidence is needed on infrastructure, access, cultural and linguistic relevance, lecturer readiness, and institutional governance, so that AI use remains ethical, fair, and educationally meaningful. Advancing this direction will support more realistic policy decisions, strengthen teaching practice, and ensure that innovation responds to African needs rather than assumptions designed elsewhere.

11 FINAL THOUGHTS

This article, drawing on established literature and grounded reasoning, indicates that large language models can enhance mathematical, statistical, and computational learning when applied responsibly, guided, and with a clear educational purpose. Evidence shows that these systems help students understand mathematical reasoning through step-by-step explanations, engage more confidently in programming tasks, strengthen computational thinking, handle statistics and data analysis more effectively, improve academic writing and reporting, and develop stronger motivation and engagement. These gains tend to be strongest when learners possess AI literacy, when prompting competence improves, and when lecturer guidance and verification remain central to learning practice. Responsible educational planning, ethical policy direction, assessment reform, and clear institutional guidance are therefore necessary to maximise benefit.

For Nigeria and the wider African region, large language models present meaningful opportunities to reduce instructional gaps, build student confidence, and increase participation in quantitative disciplines critical to national development. Although risks remain, a disciplined academic culture, structured guidance, and thoughtful integration can effectively manage them. The central challenge for institutions is to integrate these technologies responsibly while maintaining academic standards, but to apply them with fairness, wisdom, and firm commitment to learning. With careful planning, African universities can use LLMs to strengthen competence, widen opportunity, and prepare a confident generation capable of succeeding in mathematically, statistically, and computationally demanding futures.

REFERENCES

Abd-Alrazaq, A., AlSaad, R., Alhuwail, D., Ahmed, A., Healy, P. M., Latifi, S., Aziz, S., Damseh, R., Alabed Alrazak, S., & Sheikh, J. (2023). Large language models in medical education: Opportunities, challenges, and future directions. *JMIR Medical Education, 9*, Article e48291. https://doi.org/10.2196/48291 [journal article]

Aler Tubella, A., Mora-Cantallops, M., & Nieves, J. C. (2024). *How to teach responsible AI in higher education: Challenges and opportunities*. *Ethics and Information Technology, 26*, Article 3. https://doi.org/10.1007/s10676-023-09733-7 [journal article]

Andoko, B. S., Subhi, D. H., Wafiuddin, N. F., Wang, X., & Hirashima, T. (2025). Enhancing EFL reading comprehension via an AI-chatbot-guided Toulmin mapping in Viat-Map. *Utamax: Journal of Ultimate Research and Trends in Education, 7*(2), 144–158. https://doi.org/10.31849/9nwj8r52 [journal article]

Arif, H., & Naeem, J. (2025). The impact of generative AI on learner autonomy and critical thinking in English as a foreign language (EFL) writing classrooms. *Journal of Applied Linguistics and TESOL (JALT), 8*(3), 2264–2275. https://doi.org/10.63878/jalt1249 [journal article]

Bauer, E., Greiff, S., Graesser, A. C., Scheiter, K., & Sailer, M. (2025). Looking beyond the hype: Understanding the effects of AI on learning. *Educational Psychology Review, 37*, Article 45. https://doi.org/10.1007/s10648-025-10020-8 [journal article]

Canonigo, A. M. (2024). Levering AI to enhance students' conceptual understanding and confidence in mathematics. *Journal of Computer Assisted Learning, 40*(6), 3215–3229. https://doi.org/10.1111/jcal.13065 [journal article]

Chance, E. A. (2025). The combined impact of AI and VR on interdisciplinary learning and patient safety in healthcare education: A narrative review. *BMC Medical Education, 25*, Article 1039. https://doi.org/10.1186/s12909-025-07589-7 [journal article]

Chansa, T., Mumbi, M., Sain, Z. H., Pedzisai, D. R., & Mweemba, B. (2025). Exploring the use of AI-powered chatbots and writing assistants on academic integrity in Zambia's higher learning institutions. *Asian Journal of Research in Computer Science, 18*(4), 285–300. https://doi.org/10.9734/ajrcos/2025/v18i4620 [journal article]

Chen, Y. (2025). Evaluation of the impact of AI driven personalised learning platform on medical students learning performance. *Frontiers in Medicine, 12*, Article 1610012. https://doi.org/10.3389/fmed.2025.1610012 [journal article]

Choi, S., & Kim, H. (2025). The impact of a large language model-based programming learning environment on students' motivation and programming ability. *Education and Information Technologies, 30*, 8109–8138. https://doi.org/10.1007/s10639-024-13107-x [journal article]

D'Souza, R. F., Mathew, M., Mishra, V., & Surapaneni, K. M. (2024). Twelve tips for addressing ethical concerns in the implementation of artificial intelligence in medical education. *Medical Education Online, 29*(1), Article 2330250. https://doi.org/10.1080/10872981.2024.2330250 [journal article]

Dewey, J. (1938). *Logic: The Theory of Inquiry*. New York: Holt; and *Experience and Education*. New York: Macmillan. [authored book]

Dogaru, M., Pisică, O., Popa, C.-Ș., Răgman, A.-A., & Tololoi, I.-R. (2025). The perceived impact of artificial intelligence on academic learning. *Frontiers in Artificial Intelligence, 8*, Article number 1611183. https://doi.org/10.3389/frai.2025.1611183 [journal article]

Essien, A., Bukoye, O., O'Dea, C., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, 49, 865–882. https://doi.org/10.1080/03075079.2024.2316881 [journal article]

Esuong, U. U., Owan, V. J., Edoho, E. A., & Eni, B. E. (2022). Mathematics symbol instruction and senior secondary students' achievement in word problems: A quasi-experimental study. *Pedagogical Research*, *8*(1), Article ID em0142. https://doi.org/10.29333/pr/12595 [journal article]

Fang, M., Wan, X., Lu, F., Xing, F., & Zou, K. (2025). MathOdyssey: Benchmarking mathematical problem-solving skills in large language models using Odyssey Math data. *Scientific Data, 12*, Article number 1392. https://doi.org/10.1038/s41597-025-05283-3 [journal article]

Feigerlova, E., Hani, H., & Hothersall-Davies, E. (2025). A systematic review of the impact of artificial intelligence on educational outcomes in health professions education. *BMC Medical Education, 25*, Article number 129. https://doi.org/10.1186/s12909-025-06719-5 [journal article]

García-Martínez, I., Fernández-Batanero, J. M., Fernández-Cerero, J., & León, S. P. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: Systematic review and meta-analysis. *Journal of New Approaches in Educational Research, 12*, 171–197. https://doi.org/10.7821/naer.2023.1.1240 [journal article]

Granjeiro, J. M., Del Bel Cury, A. A., Cury, J. A., Bueno, M., Sousa-Neto, M. D., & Estrela, C. (2025). The future of scientific writing: AI tools, benefits, and ethical implications. *Brazilian Dental Journal, 36*, Article number e25-6471. https://doi.org/10.1590/0103-644020256471 [journal article]

Herbawani, L., Susanti, A., & Adnani, Q. (2025). The Revolution in Midwifery Education: How AI and Deep Learning are Transforming Outcome-Based Assessments?. *Advances in Medical Education and Practice*, 16, 1579–1599. https://doi.org/10.2147/amep.s543098 [journal article]

Jošt, G., Taneski, V., & Karakatič, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences, 14*(10), 4115. https://doi.org/10.3390/app14104115 [journal article]

Kiesler, N., Schiffner, D., & Schneider, J. (2025). *Enhancing computing education through large language models*. *Journal of Computer Assisted Learning, 41*(3), Article number e70048. https://doi.org/10.1111/jcal.70048 [journal article]

Kikasu, E. T., Doba, Y. G., Pillay, S. S., & Mungeni Kankisingi, G. (2025). Higher education and its contribution to economies of African countries: Move towards competence-based and skills demand-driven standards in collaboration with industry. *Journal of Education and Learning, 14*(3), 21–50. https://doi.org/10.5539/jel.v14n3p21 [journal article]

Kim, J. (2023). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies*, 29, 8693–8724. https://doi.org/10.1007/s10639-023-12109-5 [journal article]

Kumar, H., Rothschild, D. M., Goldstein, D. G., & Hofman, J. M. (2023). Math education with large language models: Peril or promise? SSRN. https://doi.org/10.2139/ssrn.4641653 [journal article]

Lademann, J., Henze, J., & Becker-Genschow, S. (2025). Augmenting learning environments using AI custom chatbots: Effects on learning performance, cognitive load, and affective variables. *Physical Review Physics Education Research, 21*, Article number 010147. https://doi.org/10.1103/PhysRevPhysEducRes.21.010147 [journal article]

Lake, B., & Baroni, M. (2023). Human-like systematic generalization through a meta-learning neural network. *Nature*, 623, 115 - 121. https://doi.org/10.1038/s41586-023-06668-3 [journal article]

Li, S., Liu, J., & Dong, Q. (2025). Generative artificial intelligence-supported programming education: Effects on learning performance, self-efficacy and processes. *Australasian Journal of Educational Technology*. *41*(3). https://doi.org/10.14742/ajet.9932 [journal article]

Liu, C., Hou, J., Tu, Y., Wang, Y., & Hwang, G. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 31, 5614–5632. https://doi.org/10.1080/10494820.2021.2012812 [journal article]

Luo, Y., & Yang, Y. (2024). Large language model and domain-specific model collaboration for smart education. *Frontiers of Information Technology & Electronic Engineering*, 25, 333–341. https://doi.org/10.1631/fitee.2300747 [journal article]

Malik, A., Khan, M., Hussain, K., Qadir, J., & Tarhini, A. (2024). AI in higher education: unveiling academicians' perspectives on teaching, research, and ethics in the age of ChatGPT. *Interactive Learning Environments*, 33, 2390–2406. https://doi.org/10.1080/10494820.2024.2409407 [journal article]

Matzakos, N., & Moundridou, M. (2025). Exploring large language models integration in higher education: A case study in a mathematics laboratory for civil engineering students. *Computer Applications in Engineering Education, 33*(3), Article number e70049. https://doi.org/10.1002/cae.70049 [journal article]

Matzakos, N., Doukakis, S., & Moundridou, M. (2023). Learning Mathematics with Large Language Models: A Comparative Study with Computer Algebra Systems and Other Tools. *Int. J. Emerg. Technol. Learn.*, 18, 51–71. https://doi.org/10.3991/ijet.v18i20.42979 [journal article]

Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia:

Opportunities and challenges. *BioData Mining, 16*, Article number 20. https://doi.org/10.1186/s13040-023-00339-9 [journal article]

Mohamed, A. M., Shaaban, T. S., Bakry, S. H., Guillén-Gámez, F. D., & Strzelecki, A. (2025). Empowering the faculty of education students: Applying AI's potential for motivating and enhancing learning. *Innovative Higher Education, 50*, 587–609. https://doi.org/10.1007/s10755-024-09747-z [journal article]

Mohammed, I. A., Owan, V. J., Thomas, G., & Ekpoto, D. F. (2025). Multi group analysis of demographic differences in higher education students ChatGPT use behaviour within a modified UTAUT2 model. *Discover Education*, *In Press*. https://doi.org/10.1007/s44217-025-01018-z [journal article]

Muringa, T. P. (2025). Exploring ethical dilemmas and institutional challenges in AI adoption: A study of South African universities. *Frontiers in Education, 10*, Article number 1628019. https://doi.org/10.3389/feduc.2025.1628019 [journal article]

Naik, A., Yin, J., Kamath, A., Q., Wu, S., Murray, R., Bogart, C., Sakr, M., & Rose, C. (2024). Providing tailored reflection instructions in collaborative learning using large language models. *Br. J. Educ. Technol.*, 56, 531–550. https://doi.org/10.1111/bjet.13548 [journal article]

Neshaei, S. P., Mejia-Domenzain, P., Davis, R. L., & Käser, T. (2025). Metacognition meets AI: Empowering reflective writing with large language models. *British Journal of Educational Technology, 56*(5), 1864–1896. https://doi.org/10.1111/bjet.13601 [journal article]

Nzoka, F. K. (2024). Reforming education in Kenya for global competitiveness: A case of competency based curriculum. *European Journal of Contemporary Education and E Learning, 2*(3), 141–150. https://doi.org/10.59324/ejceel.2024.2(3).12 [journal article]

Ofem, U. J., Owan, V. J., Iyam, M. A., Udeh, M. I., Anake, P. M., & Ovat, S. V. (2025). Students' perceptions, attitudes and utilisation of ChatGPT for academic dishonesty: Multigroup analyses via PLS–SEM. *Education and Information Technologies*, *30*(1), 159–187. https://doi.org/10.1007/s10639-024-12850-5 [journal article]

Olayinka, T., Waghid, Z., Mataka, T., & Adegoke, O. (2024). Demystifying Lesotho, Rwandan and Nigerian educators' viewpoints on smart technologies supporting AI in higher education. *Education and Information Technologies*, 29, 24285–24307. https://doi.org/10.1007/s10639-024-12820-x [journal article]

Omeh, C., Ayanwale, M., Mnguni, L., & Olelewe, C. (2025). Fostering programming skill and critical thinking through AI-assisted PBL integration. *Journal of New Approaches in Educational Research*, 14. Article number 22. https://doi.org/10.1007/s44322-025-00041-0 [journal article]

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(8), Article ID em2307. https://doi.org/10.29333/ejmste/13428 [journal article]

Owan, V. J., Chukwu, C. O., Agama, V. U., Owan, T. J., Ogar, J. O., & Etorti, I. J. (2025). Acceptance and use of artificial intelligence for self-directed research learning among postgraduate students in Nigerian public universities. *Discover Education*, *4*(1), Article No 329. https://doi.org/10.1007/s44217-025-00770-6 [journal article]

Owan, V. J., Mohammed, I. A., Bello, A., & Shittu, T. A. (2025). Higher education students' ChatGPT use behavior: Structural equation modelling of contributing factors through a modified UTAUT model. *Contemporary Educational Technology*, *17*(4), Article No. ep592. https://doi.org/10.30935/cedtech/17243 [journal article]

Park, A., & Kim, T. (2025). Code suggestions and explanations in programming learning: Use of ChatGPT and performance. *The International Journal of Management Education, 23(2), Article number 101119* https://doi.org/10.1016/j.ijme.2024.101119 [journal article]

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of Academic Ethics, 20*(2), Article number 07. https://doi.org/10.53761/1.20.02.07 [journal article]

Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: International Universities Press.

Picardal, M. T. (2025). Utilization of AI-driven smart prompts in academic research in higher education institutions. *International Journal of Learning, Teaching and Educational Research, 24*(9), 384–404. https://doi.org/10.26803/ijlter.24.9.19 [journal article]

Plaat, A., Wong, A., Verberne, S., Broekens, J., Van Stein, N., & Bäck, T. (2025). Multi-step reasoning with large language models: A survey. *ACM Computing Surveys, 58*(6), 1–35. https://doi.org/10.1145/3774896 [journal article]

Pratama, H. (2025). Training students to identify and correct fabricated references in ChatGPT-generated literature reviews. *Conference on English Language Teaching*, Article number 1335. https://doi.org/10.24090/celti.2025.1335 [journal article]

Qu, C., Dai, S., Wei, X., Cai, H., Wang, S., Yin, D., Xu, J., & Wen, J.-R. (2025). Tool learning with large language models: A survey. *Frontiers of Computer Science, 19*, Article number 198343. https://doi.org/10.1007/s11704-024-40678-2 [journal article]

Ramadana, M. R., Ulfa, S., & Soepriyanto, Y. (2025). Analyzing student cognitive engagement in AI-based learning using prompting techniques. *Indonesian Journal of Artificial Intelligence and Data Mining, 8*(2), 326–336. https://doi.org/10.24014/ijaidm.v8i2.36160 [journal article]

Raza, M., Jahangir, Z., Riaz, M. B., Saeed, M. J., & Sattar, M. A. (2025). Industrial applications of large language models. *Scientific Reports, 15*, Article number 13755. https://doi.org/10.1038/s41598-025-98483-1 [journal article]

Rojo, A., Fontaines-Ruiz, T., Bracho, A., & Rincón, L. (2025). From Digital Natives to AI Natives: Emerging Competencies and Media and Information Literacy in Higher Education. *Education Sciences*. *15*(9), Article number 1134. https://doi.org/10.3390/educsci15091134 [journal article]

Saal, P., Chetty, K., Ntshayintshayi, N., Moosa, T., & Masuku, N. (2025). A scoping review of the integration of artificial intelligence in primary and secondary schools from 2020 to 2024: Policy implications for South Africa. *Journal of Education*. *(98)*, Article i98a04 https://doi.org/10.17159/2520-9868/i98a04 [journal article]

Safranek, C., Sidamon-Eristoff, A., Gilson, A., & Chartash, D. (2023). The Role of Large Language Models in Medical Education: Applications and Implications. *JMIR Medical Education*, 9, Article e50945. https://doi.org/10.2196/50945 [journal article]

Salhab, R. (2024). AI Literacy across Curriculum Design: Investigating College Instructor's Perspectives. *Online Learning, 28(2), Article 4426.* https://doi.org/10.24059/olj.v28i2.4426 [journal article]

Shang, X., Li, R., & Li, Y. (2023). The effect of structured stepwise presentations on students' fraction learning: An eye-tracking study. *Frontiers in Psychology, 14*, Article 1125589. https://doi.org/10.3389/fpsyg.2023.1125589 [journal article]

Shelton, C. (2016). Students who developed logical reasoning skills reported improved confidence in drug dose calculation: Feedback from remedial maths classes. *Nurse education today*, 41, 6–11. https://doi.org/10.1016/j.nedt.2016.03.007 [journal article]

Shen, M., Shen, Y., Liu, F., & Jin, J. (2025). Prompts, privacy, and personalized learning: Integrating AI into nursing education, A qualitative study. *BMC Nursing, 24*, Article number 470. https://doi.org/10.1186/s12912-025-03115-8 [journal article]

Shi, H., Xu, Z., Wang, H., Qin, W., Wang, W., Wang, Y., Wang, Z., Ebrahimi, S., & Wang, H. (2025). Continual learning of large language models: A comprehensive survey. *ACM Computing Surveys, 58*(5), Article 120. https://doi.org/10.1145/3735633 [journal article]

Song, N. (2024). Higher education crisis: Academic misconduct with generative AI. *Journal of Contingencies and Crisis Management, 32*(1), Article e12532. https://doi.org/10.1111/1468-5973.12532 [preprint]

Song, X., Lee, L., Xie, K., Liu, X., Deng, X., & Hong, Y. (2025). StatLLM: A dataset for evaluating the performance of large language models in statistical analysis [Preprint]. arXiv. https://arxiv.org/abs/2502.17657 [journal article]

Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artif9cial Intelligence*, 4, Article 100127. https://doi.org/10.1016/j.caeai.2023.100127 [journal article]

Sweller, J., van Merrienboer, J.J.G. & Paas, F.G.W.C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*, 251–296. https://doi.org/10.1023/A:1022193728205 [journal article]

Tran, M., Balasooriya, C., Jonnagaddala, J., Leung, G. K.-K., Mahboobani, N., Ramani, S., Rhee, J., Schuwirth, L., Najafzadeh-Tabrizi, N. S., Semmler, C., & Wong, Z. S. Y. (2025). Situating governance and regulatory concerns for generative artificial intelligence and large language models in medical education. *NPJ Digital Medicine, 8*, Article 315. https://doi.org/10.1038/s41746-025-01721-z [journal article]

Tyndall, E., Gayheart, C., Some, A., Genz, J., Wagner, T., & Langhals, B. (2025). Impact of retrieval augmented generation and large language model complexity on undergraduate exams created and taken by AI agents. *Data & Policy, 7*, Article e57. https://doi.org/10.1017/dap.2025.10024 [journal article]

Wang, M., Zhu, J., Hwang, G.-J., Chang, S.-C., Yang, Q.-F., & Zhang, D. (2025). Boosting student engagement in STEM: Integrating large language model-based virtual agents into alternate reality games. *Journal of Computer Assisted Learning, 41*(6), Article e70139. https://doi.org/10.1111/jcal.70139 [journal article]

Wang, T. (2019). Competence for students' future: Curriculum change and policy redesign in China. *ECNU Review of Education, 2*(2), 234–245. https://doi.org/10.1177/2096531119850905 [journal article]

Webb, T., Holyoak, K., & Lu, H. (2022). Emergent analogical reasoning in large language models. *Nature Human Behaviour*, 7, 1526 - 1541. https://doi.org/10.1038/s41562-023-01659-w [journal article]

Yan, L., Sha, L., Zhao, L., Li, Y., Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55, 90–112. https://doi.org/10.1111/bjet.13370 [journal article]

Yang, M., Jiang, S., Li, B., Herman, K., Luo, T., Moots, S. C., & Lovett, N. (2025). Analysing nontraditional students' ChatGPT interaction, engagement, self-efficacy and performance: A mixed-methods approach. *British Journal of Educational Technology, 56*(5), 1973–2000. https://doi.org/10.1111/bjet.13588 [journal article]

Ye, B., Xi, Y., & Zhao, Q. (2024). Optimising mathematical problem-solving reasoning chains and personalized explanations using large language models: A study in applied mathematics education. *Journal of AI-Powered Medical Innovations, 3*(1). https://doi.org/10.60087/Japmi.Vol.03.Issue.01.Id.005 [journal article]

Yin, S., Fu, C., Zhao, S., Li, K., Sun, X., Xu, T., & Chen, E. (2024). A survey on multimodal large language models. *National Science Review, 11*(12), Article nwae403. https://doi.org/10.1093/nsr/nwae403 [journal article]

Yousef, M., Mohamed, K., Medhat, W., Mohamed, E., Khoriba, G., & Arafa, T. (2024). BeGrading: large language models for enhanced feedback in programming education. *Neural Computing and Applications*, 37, 1027–1040. https://doi.org/10.1007/s00521-024-10449-y [journal article]

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21, 1–29. https://doi.org/10.1186/s41239-024-00453-6 [journal article]

Zdravkova, K., & Ilijoski, B. (2025). The impact of large language models on computer science student writing. *International Journal of Educational Technology in Higher Education, 22*, Article 32. https://doi.org/10.1186/s41239-025-00525-1 [journal article]

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*, Article 28. https://doi.org/10.1186/s40561-024-00316-7 [journal article]

Zhang, Q., Huang, Z., Huang, Y., Wang, G., Zhang, R., Yang, J., Cheng, Y., Chen, B., Wang, H., Qiu, K., & Chen, H. (2025). Generative AI in medical education: Feasibility and educational value of LLM-generated clinical cases with MCQs. *BMC Medical Education, 25*, Article 1502. https://doi.org/10.1186/s12909-025-08085-8 [journal article]

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon, 10*(3), Article e25370. https://doi.org/10.1016/j.heliyon.2024.e25370 [journal article]

Zhong, Q., Wang, K., Xu, Z., Ding, L., Liu, J., & Du, B. (2026). Achieving >97% on GSM8K: Deeply understanding the problems makes LLMs better solvers for math word problems. *Frontiers of Computer Science, 20*, Article 2001310. https://doi.org/10.1007/s11704-025-41102-z [journal article]

Zhou, Z. (2025). Large language models and their evolution. *Applied and Computational Engineering, 173*. https://doi.org/10.54254/2755-2721/2025.25586 [journal article]

Zhui, L., Yhap, N., Liping, L., Zhengjie, W., Zhonghao, X., Xiaoshu, Y., Hong, C., Xuexiu, L., & Wei, R. (2024). Impact of large language models on medical education and teaching adaptations. *JMIR Medical Informatics, 12*, Article e55933. https://doi.org/10.2196/55933 [journal article]