Generative Artificial Intelligence and Assessment Integrity in Nigerian Universities: Reimagining Educational Measurement and Evaluation in the Age of Intelligent Technologies

Iwunze Jacinta Chika (Ph.D)

Federal Ministry of Education, Federal Government College, Enugu, Nigeria.

DOI: https://doi.org/10.68050/JAMS.2026.319

Received: 31 Aug 2026; Accepted: 05 Sept 2026; Published: 14 November 2026

ABSTRACT

Higher education has significantly been transformed by the rapid emergence of Generative Artificial Intelligence (GenAI) through its introduction of new possibilities for teaching, learning, research, and assessment. The increasing accessibility of GenAI has however generated critical concerns regarding assessment integrity, particularly the validity, authenticity, reliability, fairness, and transparency of assessment outcomes. The implications of GenAI for Educational Measurement and Evaluation is examined in this conceptual review, with particular emphasis on Nigerian universities. The article draws on Technology Acceptance Model and Messick's Unified Theory of Validity, and argues that the central challenge posed by GenAI is not the technology itself but the capacity of existing assessment systems to continue producing valid and defensible evidence of student learning. Contemporary scholarship on AI adoption, assessment transformation, academic integrity, authentic assessment, AI literacy, and institutional governance are synthesized in the review. It advocates a shift from detection-oriented approaches towards GenAI-responsive assessment practices that emphasise higher-order thinking, authentic performance, transparency, and responsible technology integration. In conclusion, the article is of the view that to preserve educational quality while harnessing the opportunities offered by intelligent technologies, Nigerian universities require context-sensitive policies, strengthened digital capacity, lecturer professional development, and innovative assessment redesign.

Keywords: Generative Artificial Intelligence; Assessment Integrity; Educational Measurement; Higher Education; Ai Literacy, Nigerian

UNIVERSITIES

Disclosure: OpenAI's ChatGPT was used during the preparation of this manuscript to assist with language editing, restructuring, and improvement of clarity and academic expression. All AI-assisted output was critically reviewed and revised by the author. The author independently verified the factual content and references and assumes full responsibility for the research design, analysis, interpretation, arguments, and conclusions presented in the article.

1. INTRODUCTION

The rapid advances in Artificial Intelligence (AI) are profoundly transforming the contemporary landscape of higher education. Particular mention is made in this regard of Generative Artificial Intelligence (GenAI). While earlier educational technologies primarily supported information retrieval, communication, or administrative automation, GenAI systems possess the capacity to generate human-like text, images, code, explanations, and other forms of digital content. ChatGPT and other emerging large language models have therefore introduced unprecedented possibilities for supporting teaching, learning, research, and academic productivity (Kasneci et al., 2023; UNESCO, 2023).

GenAI applications have expanded rapidly within higher education. AI tools are increasingly being employed

by students for brainstorming, summarizing academic materials, improving writing, translating texts, solving problems, and supporting research activities. Lecturers have similarly begun exploring GenAI for lesson preparation, instructional design, feedback generation, and assessment development. The potential of AI to enhance educational efficiency, personalize learning experiences, and promote wider access to academic resources are demonstrated by these applications (Baidoo-Anu & Ansah, 2023; Holmes et al., 2022).

However, concerns have simultaneously been intensified by the growing adoption of GenAI regarding the integrity of university assessment systems. The role of Educational assessment in determining student achievement, certifying competence, informing educational decisions, and maintaining public confidence in academic qualifications, is fundamental. Assessment systems must, as a result, provide credible evidence that students possess the knowledge, skills, and competencies represented by their assessment scores (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). Traditional assumptions underlying many assessment practices are challenged by the capacity of GenAI to generate sophisticated academic responses. It is often assumed that submitted works as responses to conventional assignments, essays, take-home examinations, and written reports primarily represent students' independent intellectual effort. This assumption is complicated by the availability of AI systems capable of producing coherent and academically acceptable responses, thereby raising fundamental questions concerning authorship, authenticity, and the interpretation of assessment outcomes (Cotton et al., 2024; Perkins, 2023).

The central concern from the perspective of Educational Measurement and Evaluation is not merely whether students use AI but whether assessment scores continue to represent the intended constructs being measured. The unified theory of validity by Messick (1989) emphasizes that validity concerns the meaning and consequences of interpretations derived from assessment scores. If assessment outcomes increasingly reflect students' ability to access and manipulate AI-generated content rather than their own knowledge, reasoning, and competence, the validity of those interpretations may therefore, be compromised.

Within Nigerian universities, where higher education institutions are simultaneously pursuing digital transformation while facing challenges associated with technological infrastructure, unequal access, institutional preparedness, and quality assurance, these concerns are particularly significant. There remains a relatively limited empirical and conceptual discussion addressing the implications of GenAI for assessment integrity, despite the fact that Nigerian universities increasingly encounter AI-enabled educational practices.

GenAI adoption, academic integrity concerns, educational opportunities, and assessment redesign have been predominantly examined by existing international scholarships (Kasneci et al., 2023; Xia et al., 2024). There remains, however, a need for context-sensitive analysis examining how developing higher education systems, such as Nigeria's, can respond to GenAI while maintaining assessment credibility and educational standards.

The implications of Generative Artificial Intelligence for assessment integrity within Nigerian universities are thus examined in this article. It entails exploring the relationship between AI adoption, assessment validity, institutional preparedness, and responsible technology integration. The article strongly presents the argument that sustainable responses to GenAI require a transition from restrictive, detection-focused approaches towards assessment systems designed around authenticity, higher-order thinking, transparency, and defensible evidence of learning.

The article tends to contribute to Educational Measurement and Evaluation scholarship by providing a conceptual framework for understanding assessment integrity in AI-enabled educational environments. It proposes, through the integration of the Technology Acceptance Model (Davis, 1989) and Messick's Unified Theory of Validity, that the future of assessment depends not on resisting technological change but on redesigning assessment practices capable of preserving validity and public trust in an era of intelligent technologies(Messick, 1989)

2. CONCEPTUAL REVIEW METHODOLOGY

The implications of Generative Artificial Intelligence (GenAI) for assessment integrity in Nigerian universities

are examined in this article, adopting a structured conceptual-review approach. Rather than to estimate effect sizes or conducts a statistical Meta analysis, the review was designed to integrate theoretical perspectives, empirical findings and relevant policy literature. Thus the methodological procedure was organized around a transparent search strategy, explicit selection criteria and thematic synthesis of the literature.

2.1 Strategy for Literature Search

Major scholarly databases and academic search platforms, including Scopus, Web of Science, ERIC, Science Direct and Google Scholar were the sources of literature relevant to the research. The principal publication period stretched from January 2021 to December 2025, reflecting the rapid expansion of scholarship on Generative Artificial Intelligence and higher education following the emergence of publicly accessible large language models. Where earlier publications provided seminal theoretical, measurement or policy foundations for the analysis, they were retained. Davis (1989) on the Technology Acceptance Model, Fishbein and Ajzen (1975) on the Theory of Reasoned Action, Messick (1989) on validity, and AERA, APA and NCME (2014) on educational and psychological testing, all represent such.

The search involved the combination of terms relating to the technology, educational context and assessment problem. "Generative artificial intelligence", "GenAI", "ChatGPT", "large language models", "higher education", "universities", "assessment", "academic integrity", "assessment validity", "authentic assessment", "reliability", "fairness", "transparency", "AI literacy", "institutional preparedness", "assessment redesign", "Nigeria" and "Africa" were the principal search terms.

Boolean combinations were useful in connecting the principal concepts. The following constituted a representative search string:

("generative artificial intelligence" OR GenAI OR ChatGPT OR "large language models") AND ("higher education" OR university*) AND (assessment OR "academic integrity" OR validity OR authenticity OR reliability OR fairness) AND (Nigeria OR Africa).

To capture literature dealing specifically with AI literacy, institutional governance, assessment transformation and responsible AI utilization, additional combinations, were used where necessary.

2.2 Inclusion and Exclusion Criteria

The underlisted categories of literature were considered relevant for inclusion-those that

-focused on higher education, universities, teaching and learning, or university assessment;

-examined Generative AI, large language models or closely related AI applications with implications for educational assessment;

-addressed one or more of the following issues: academic integrity, validity, authenticity, reliability, fairness, transparency, AI literacy, institutional preparedness or assessment redesign; -- provided empirical findings, theoretical analysis, systematic or conceptual synthesis, or authoritative policy guidance;

- was published primarily between 2021 and 2025, except for seminal theoretical and measurement sources; and

- was sufficiently relevant to the Nigerian or broader African higher-education context, or provided evidence transferable to that context.

Studies were excluded where they focused exclusively on primary or secondary education, technical development of AI without educational implications, non-educational applications of AI, or issues having no substantive relationship with assessment, academic integrity, educational quality or institutional governance. Duplicate records and publications that only mentioned AI without substantive engagement with the review problem were also excluded.

2.3 Literature Synthesis

The selected literature was synthesized thematically. Statistical aggregation was considered inappropriate for the purpose of this article, because the reviewed publications comprised heterogeneous empirical, theoretical, conceptual and policy contributions. Recurring concepts, relationships, challenges and proposed responses were identified and organized into the following six principal analytical themes, instead:

Generative AI adoption and educational use; assessment integrity and academic misconduct;

Assessment validity and construct representation; authentic and process-oriented assessment;

AI literacy and institutional preparedness; and institutional governance, fairness and transparency.

The themes were subsequently interpreted through the Technology Acceptance Model (Davis, 1989) and Messick's Unified Theory of Validity (Messick, 1989). TAM was used to explain the conditions influencing stakeholder acceptance and utilisation of GenAI, while Messick's framework was used to examine whether assessment evidence generated in AI-enabled environments remains meaningful, valid and defensible.

The synthesis consequently moved from descriptive identification of issues to conceptual integration. Particular attention was given to relationships among institutional preparedness, AI literacy, perceived usefulness, perceived ease of use, responsible GenAI utilization, assessment transformation and assessment-integrity outcomes. This process provided the basis for the explicit conceptual framework developed in Section 5.

The methodological orientation is consistent with the purpose of a conceptual review: to organize dispersed scholarship around a clearly defined problem, identify gaps in existing explanations and develop a theoretically informed framework capable of guiding subsequent empirical investigation.

3. CONCEPTUAL FRAMEWORK

The Technology Acceptance Model (TAM) is integrated with Messick's Unified Theory of Validity by the conceptual framework proposed for this article, to explain how Generative Artificial Intelligence may influence assessment integrity in Nigerian universities. The framework views an assessment integrity beyond simply being the prevention of examination malpractice or plagiarism. Instead, it conceptualized assessment integrity as the capacity of an assessment system to generate valid, authentic, reliable, fair, transparent and defensible evidence of student learning under conditions in which GenAI is increasingly available.

A directional relationship among five major components: (1) antecedent conditions; (2) responsible GenAI utilization; (3) assessment transformation; (4) validity-assurance mechanisms; and (5) assessment-integrity outcomes, proposed by the framework.

3.1 Antecedent Conditions

The conditions that shape stakeholders' capacity and willingness to engage with GenAI constituted the first component. These include:

AI literacy among lecturers and students; institutional policies and governance structures; technological infrastructure; perceived usefulness of GenAI; and perceived ease of use.

Whereas, perceived usefulness and perceived ease of use are derived principally from TAM (Davis, 1989), AI literacy, institutional preparedness and technological infrastructure represent contextual conditions necessary for responsible adoption.

Thus, the framework proposes that the likelihood of GenAI adoption and meaningful utilization are increased

by stronger AI literacy, institutional preparedness and technological capacity, together with greater perceived usefulness and ease of use, increase the likelihood of GenAI adoption and meaningful utilization.

3.2 Responsible GenAI Utilization

Responsible GenAI utilization is the second component. It refers to the ethical, transparent and educationally purposeful use of GenAI for activities such as:

Learning support; research assistance; instructional development; formative feedback; brainstorming and idea generation; and assessment enhancement.

Responsible utilization is distinguished by the framework from undisclosed substitution of students' intellectual effort. The importance of the distinction is owing to the fact that GenAI adoption in itself is neither inherently beneficial nor inherently harmful. Its educational consequences depend substantially on how the technology is used and governed.

3.3 Assessment Transformation

Pressure is created by Responsible GenAI utilization for universities to reconsider assessment design. Thus, the third component concerns assessment transformation through:

Authentic assessment; performance-based tasks; process-oriented evaluation; reflective activities; oral defence; iterative submissions; application-based tasks; and opportunities for students to demonstrate ownership of their work.

Assessment systems relying heavily on final written products are more vulnerable to substitution by AI-generated content. That is the underlying proposition. Richer evidence of actual competence can be provided by assessment systems that capture reasoning, decision-making, application, reflection and performance.

3.4 Validity-Assurance Mechanisms

The measurement bridge between assessment transformation and assessment-integrity outcomes is constituted by Validity assurance. Validity -assurance is a component that draws directly on Messick's (1989) Unified Theory of Validity. It involves continuous examination of:

Alignment between assessment tasks and intended learning outcomes; construct representation;

construct-irrelevant variance; scoring consistency; fairness of assessment conditions;

transparency of AI-use expectations; human oversight; and consequences of assessment decisions.

The framework proposes that assessment transformation should not automatically be assumed to guarantee assessment integrity. Instead, redesigned assessment practices must themselves be evaluated to determine whether the evidence they generate supports defensible interpretations of student competence.

3.5 Assessment-Integrity Outcomes

Assessment integrity as the final component is expressed through: validity; authenticity; reliability; fairness; transparency; and defensible interpretation of assessment scores.

Thus, Assessment integrity is conceptualized as the ultimate outcome of the interaction among technological adoption, responsible AI use, institutional preparedness, assessment transformation and continuous validity assurance.

3.6 Directional Relationships within the Framework

The relationships proposed can be expressed as follows:

Antecedent Conditions

AI literacy + Institutional preparedness + Technological infrastructure + Perceived usefulness + Perceived ease of use

↓

Responsible GenAI Utilization

Ethical use + Transparency + Appropriate disclosure + educationally purposeful use

↓

Assessment Transformation

Authentic assessment + Performance tasks + Process evidence + Reflection + Oral defence + Application

↓

Validity Assurance

Construct alignment + Construct-irrelevant variance control + Reliability + Fairness + Transparency + Human oversight

↓

Assessment Integrity Outcomes

Validity + Authenticity + Reliability + Fairness + Transparency + Defensible assessment decisions

A feedback relationship is also recognized by the framework. Evidence requiring universities to revise policies, strengthen AI literacy, improve infrastructure and redesign assessment practices can be generated by assessment outcomes and institutional experience. The framework should thus be understood as dynamic rather than strictly linear.

3.7 Integrated Theoretical Logic of the Framework

Two complementary explanatory dimensions are produced by the integration of TAM and Messick's framework. Should TAM explains why lecturers and students are likely to accept and utilize GenAI, particularly where they perceive the technology as useful and easy to use, Messick's framework, in contrast, asks whether the consequences of that utilization permit assessment scores to retain their intended meaning.

The framework consequently makes an important distinction between technology acceptance and assessment quality. Weak assessment integrity may co-exist with high levels of GenAI adoption of institutional governance; AI literacy and assessment redesign are inadequate. Conversely, responsible adoption supported by appropriate assessment transformation and validity assurance may allow Universities may conversely be allowed by responsible adoption supported by appropriate assessment transformation and validity assurance to benefit from GenAI, while preserving the credibility of assessment outcomes.

The frame therefore centrally proposes that:

Positive contribution of GenAI adoption to educational assessment is only when technological acceptance is

mediated by responsible utilization, supported by institutional preparedness and translated through assessment transformation and continuous validity assurance into credible evidence of student learning.

3.8 Propositions Derived from the Conceptual Framework

The following propositions are generated by the framework for subsequent empirical investigation:

P1: Greater capacity for responsible GenAI utilization is associated with higher levels of AI literacy and institutional preparedness.

P2: Perceived usefulness and perceived ease of use influence GenAI adoption positively, consistent with the Technology Acceptance Model.

P3: Assessment integrity is not guaranteed by GenAI adoption alone; the consequences of GenAI adoption depend on the extent to which its utilization is responsible, transparent and institutionally governed.

P4: The evidence available for judging students' actual competence is strengthened by assessment transformation through authentic, performance-based and process-oriented tasks .

P5: Threats arising from construct-irrelevant variance, unequal technological access and inappropriate AI use are reduced by Validity-assurance mechanisms.

P6: Where institutional preparedness, AI literacy, responsible GenAI utilization, assessment transformation and validity assurance operate together, assessment integrity is strongest.

4. CONCEPTUAL REVIEW

4.1 Generative Artificial Intelligence in Higher Education

This represents a significant advancement in artificial intelligence because of its ability to create new content rather than merely analyze or classify existing information. GenAI systems, powered largely by deep learning architectures and large language models(LLMS), can generate written texts, computer programmes, images, summaries, explanations, and other forms of knowledge representation in response to user prompts (Kasneci et al., 2023)? Discussions concerning the educational applications and implications of the emergence of publicly accessible GenAI tools have been accelerated.

GenAI has been increasingly recognized as both a transformative educational resource and a disruptive technological development in higher education. Supporting personalized learning, generating instructional materials, providing automated feedback, assisting academic writing, facilitating research activities, and enhancing accessibility for diverse learners, are all included in its applications(Baidoo-Anu & Ansah, 2023; UNESCO, 2023). Opportunities for universities to improve learning efficiency and develop innovative pedagogical practices have been created by these capabilities.

The educational value of GenAI, however, depends significantly on how it is integrated into institutional practices. While knowledge construction and learner autonomy, can be supported by AI tools, inappropriate dependence on AI-generated outputs may reduce students' engagement with critical thinking, independent reasoning, and authentic knowledge production. The importance of AI literacy, ethical guidance, and institutional governance are thus increasingly emphasized by scholars as essential conditions for responsible adoption (UNESCO, 2023).

The debate surrounding GenAI in higher education has therefore shifted from questions of whether AI should be permitted towards determining how educational institutions can establish appropriate frameworks for its responsible use. It has gone beyond those in the shift, to recognize that GenAI is likely to remain an enduring component of the academic environment and that universities must develop strategies for adapting educational practices accordingly.

4.2 Assessment Integrity in Higher Education

One fundamental principle within Educational Measurement and Evaluation is Assessment integrity. This is because assessment outcomes influence certification, progression, accountability, and public confidence in educational institutions. Assessment integrity represents a broader concept involving the credibility, fairness, transparency, and defensibility of assessment processes and decisions, although it is often associated with preventing examination malpractice and plagiarism.

Assessment quality, according to the Standards for Educational and Psychological Testing (AERA et al., 2014), depends upon the extent to which scores provide appropriate evidence for intended interpretations and uses. Assessment integrity therefore requires that assessment tasks accurately represent learning outcomes, scoring procedures are consistent and fair, and assessment results support valid educational decisions.

The meaning and complexity of assessment integrity has been expanded by the emergence of GenAI. Preventing unauthorized assistance, plagiarism, and dishonest behaviour has largely been the focus of traditional approaches to academic integrity . GenAI, however, introduces a more complex situation because AI assistance may range from legitimate educational support to inappropriate substitution of students' intellectual effort. Consequently, it has become increasingly difficult to establish a distinction between acceptable assistance and academic misconduct.

Language models, as Perkins (2023) argues, require universities to reconsider conventional approaches to academic integrity because existing policies were developed before AI systems became capable of producing sophisticated academic outputs. Cotton et al. (2024), similarly contend that institutions must move beyond punitive approaches and develop clearer frameworks defining responsible AI use within assessment contexts. Assessment integrity must therefore be understood dynamically, within this changing environment. Not only does it require mechanisms for preventing misconduct, it also requires assessment designs capable of ensuring that students demonstrate genuine understanding, reasoning, creativity, and competence.

4.3 Generative Artificial Intelligence and Assessment Validity

The central concept underpinning educational assessment remains Validity. Validity is conceptualized by Messick (1989)) as an integrated judgement regarding the appropriateness and meaningfulness of interpretations derived from assessment scores. It can be seen from this perspective, that assessment validity is not an inherent property of an instrument alone but a function of how assessment evidence is interpreted and used.

GenAI comes with new challenges to validity because it may introduce sources of variance unrelated to the intended learning construct. The example can be cited of an essay assessment designed to measure students' writing ability, analytical reasoning, or conceptual understanding that may produce misleading evidence if a substantial portion of the submitted response is generated by AI without disclosure. The assessment may in such circumstances measure students' ability to prompt and refine AI outputs rather than their independent competence.

Assessment developers as emphasized by the Standards for Educational and Psychological Testing (AERA et al., 2014), must consider sources of construct-irrelevant variance that threaten score interpretation. This, in AI-enabled learning environments, requires renewed attention to assessment design, administration conditions, scoring procedures, and the relationship between assessment tasks and intended learning outcomes.

The challenge of GenAI is therefore fundamentally a challenge of maintaining validity. It is imperative for universities to ensure that assessment practices continue to provide trustworthy evidence of student learning despite changing technological conditions.

4.4 Authentic Assessment as a Response to Generative Artificial Intelligence

Authentic assessment, as a possible response to challenges created by GenAI, has gained renewed attention.

Unlike traditional assessment approaches that frequently emphasize reproduction of information, it requires learners to apply knowledge, solve realistic problems, demonstrate professional competencies, and engage in meaningful intellectual processes.

Portfolios, projects, case studies, presentations, reflective activities, simulations, collaborative tasks, and performance-based assessments, are all types of authentic assessment approaches. They are particularly relevant in AI-enabled environments because they evaluate processes of thinking, decision-making, application, and reflection rather than only final written products.

Traditional assessment methods are not meant to be abandoned due to the increasing importance of authentic assessment. The latter, rather, suggests and even underscores the need for balanced assessment systems that combine different forms of evidence to produce more comprehensive interpretations of student competence. Oral explanations, iterative submissions, drafts, reflective commentary, and opportunities for students to demonstrate ownership of their learning may all be included in such systems.

Authentic assessment, from an Educational Measurement perspective, aligns with the principle that assessment should generate meaningful evidence of the knowledge and abilities it intends to measure. Authenticity in the era of GenAI, becomes increasingly important. This is because universities must distinguish between technological assistance that enhances learning and technological substitution that obscures actual competence.

4.5 AI Literacy and Institutional Preparedness

Integrating GenAI responsibly requires more than access to technological tools. AI literacy among lecturers, students, administrators, and policymakers is exigent. AI literacy implies understanding how AI systems function, recognizing their limitations, evaluating their outputs critically, and applying them ethically within specific contexts. AI literacy, for lecturers, includes the ability to redesign assessment tasks, establish appropriate AI-use expectations, evaluate AI-assisted student work, and incorporate AI into teaching practices responsibly. AI literacy, for students, involves understanding appropriate AI applications, acknowledging AI assistance transparently, and maintaining responsibility for academic work.

Equally important is institutional preparedness. Universities require policies, infrastructure, professional development programmes, and governance mechanisms that guide responsible AI adoption. Educational institutions, UNESCO (2023), should develop human-centred AI approaches that promote inclusion, ethical responsibility, and educational quality.

Institutional preparedness, for Nigerian universities, is particularly significant because AI integration occurs within a context characterized by varying levels of technological capacity.

Strategies that address digital inequality, infrastructure limitations, staff competence, and institutional policy development are therefore required for effective implementation.

5. THEORETICAL FRAMEWORK

Two complementary theoretical perspectives: the Technology Acceptance Model (TAM) developed by Davis (1989) and Messick's Unified Theory of Validity (Messick, 1989), provide the anchor for this article. Integrating these frameworks provides a comprehensive basis for examining both the adoption of Generative Artificial Intelligence (GenAI) and its implications for assessment integrity within Nigerian universities.

TAM explains the factors influencing academic staff' and students' acceptance and utilization of GenAI technologies, while Messick's validity framework provides the measurement perspective necessary for evaluating whether assessment outcomes remain meaningful, accurate, and defensible in AI-enabled educational environments. The two theories provide together a balanced explanation of the technological and measurement dimensions of GenAI adoption in higher education.

5.1 Technology Acceptance Model (TAM)

This model was introduced by Davis (1989) as a theoretical framework for explaining users' acceptance of information technologies. TAM, derived from the Theory of Reasoned Action developed by Fishbein and Ajzen (1975), proposes that individuals' willingness to adopt a technology is primarily determined by two major beliefs: perceived usefulness and perceived ease of use.

Perceived usefulness implies the extent to which an individual believes that using a particular technology will enhance performance, productivity, or effectiveness. The degree to which an individual perceives that using the technology will require minimal effort, explains perceived ease of use. These beliefs, according to Davis (1989) influence users' attitudes, behavioural intentions, and eventual technology adoption.

TAM has been widely applied within higher education, to explain the acceptance of digital learning platforms, online learning systems, mobile technologies, and emerging educational innovations. The fact that lecturers and students' willingness to adopt AI tools is influenced by their perceptions of usefulness, accessibility, reliability, and institutional support accounts for the increase its relevance with the emergence of GenAI.

TAM, in the context of Nigerian universities, provides a useful explanation of why stakeholders may adopt or resist GenAI in teaching, learning, research, and assessment. It is more likely that academic staff who perceive GenAI as beneficial for instructional planning, feedback provision, and assessment development would integrate it into their professional practices. Students who perceive AI tools as useful for learning support, academic writing, and research assistance are more likely to employ them- similarly.

TAM alone, however, does not adequately explain whether the use of GenAI improves or threatens educational quality. While a technology may be widely accepted it may simultaneously be creating challenges for assessment validity and academic integrity. A measurement theory is therefore required to examine the consequences of GenAI adoption for assessment outcomes.

5.2 Messick's Unified Theory of Validity

This theory represents one of the most influential perspectives in modern Educational Measurement and Evaluation. The view that validity consists of separate categories such as content validity, criterion-related validity, and construct validity, was rejected by Messick. He proposed validity, instead, as a unified concept concerned with the extent to which evidence supports the interpretations and uses of assessment scores.

Validity, according to Messick(1989) involves multiple sources of evidence, including content relevance, substantive processes, structural relationships, generalisability, external relationships, and consequences of assessment use. Assessment quality is highlighted by this perspective as depending, not only on technical measurement properties, but also on the meaning and implications of assessment decisions.

The relevance of Messick's framework to GenAI can be found in its emphasis on construct representation and construct-irrelevant variance. If students' scores reflect their ability to access and manipulate AI-generated content rather than the knowledge, skills, and competencies intended to be measured, assessment outcomes in AI-enabled educational environments may be distorted.

If, for example, a university uses essay assignments to assess critical thinking and written communication, but students submit predominantly AI-generated responses without meaningful intellectual engagement, the resulting scores may no longer represent students' actual abilities. Such circumstances threaten the validity of assessment interpretations and weaken confidence in educational decisions.

An essential basis for examining how universities can preserve assessment integrity through appropriate assessment design, transparent AI-use policies, and evaluation practices capable of generating trustworthy evidence of student learning, is thus provided by Messick's framework.

5.3 Integration of TAM and Messick's Validity Framework

Integrating TAM and Messick's Unified Theory of Validity provides a comprehensive conceptual foundation for understanding GenAI and assessment integrity. By addressing why academic staff and students choose to adopt AI technologies, TAM explains the acceptance dimension of GenAI. It highlights the imperative of perceived usefulness, perceived ease of use, digital competence, and institutional support in shaping AI adoption behaviours.

The measurement dimension is addressed by Messick's framework by examining whether assessment outcomes remain valid and meaningful when AI technologies become part of the educational environment. It offers a shift of attention from technology adoption alone to the quality and consequences of assessment decisions.

These frameworks together, suggest that successful GenAI integration in universities requires two complementary conditions:

Responsible technological adoption, supported by AI literacy, institutional readiness, and ethical guidance; and

Assessment validity preservation, achieved through authentic assessment design, transparent practices, and continuous evaluation of assessment consequences.

The central argument of this article is therefore, that the future of educational assessment does not depend on preventing access to GenAI but on ensuring that assessment systems evolve sufficiently to preserve the validity, authenticity, reliability, fairness, and transparency of educational decisions.

6. EMPIRICAL REVIEW

6.1 Generative Artificial Intelligence and Higher Education Assessment: International Evidence

There has been a rapid expansion of empirical research on Generative Artificial Intelligence (GenAI) in higher education, since the emergence of large language models capable of producing human-like academic responses. Students' and lecturers' perceptions, patterns of AI adoption, academic integrity concerns, assessment transformation, and institutional responses, have all been examined in existing studies.

The educational implications of large language models was examined by Kasneci et al. (2023) who argued that GenAI presents significant opportunities for improving personalized learning, providing feedback, supporting writing development, and enhancing accessibility. The authors, however, also identified major challenges relating to inaccurate AI outputs, overdependence on AI systems, assessment validity, and the changing nature of academic work. Concluding, they opined that educational institutions must develop responsible frameworks that maximize AI benefits while reducing potential harms.

Baidoo-Anu and Ansah (2023), similarly explored the potential applications of ChatGPT in education and reported that generative AI could support teaching and learning through personalized assistance, content generation, and enhanced student engagement. They emphasized, nevertheless, that effective implementation requires careful consideration of ethical issues, digital competence, and academic integrity.

Assessment has specifically been the focus of a growing body of empirical scholarship.

The implications of ChatGPT for academic integrity was investigated by Cotton et al.(2024), who argued that higher education institutions face challenges in maintaining confidence in traditional assessment approaches. Relying on their findings, universities need to move beyond approaches based solely on detecting misconduct and instead redesign assessment practices to emphasize authentic demonstrations of learning.

Through a systematic review of research on artificial intelligence in higher education assessment, Xia et Al.(2024), found that AI is transforming assessment practices through automated feedback, adaptive

assessment, and new approaches to evaluating student learning. The review, however, also identified persistent concerns regarding fairness, transparency, privacy, validity, and the need for institutional governance. Future assessment systems, as recommended by the authors, should integrate technological innovation with human-centred educational principles.

Put together, these international studies demonstrate that GenAI is not simply a technological innovation but a major educational development requiring reconsideration of assessment purposes, methods, and quality assurance mechanisms.

6. 2 Artificial Intelligence, Assessment Integrity, and Educational Measurement

Viewed from an Educational Measurement perspective, empirical studies increasingly indicate that GenAI challenges traditional assumptions concerning assessment validity. Whether assessment scores continue to provide accurate evidence of students' knowledge and abilities when AI tools participate in the production of academic outputs is the central issue. Valid assessment, as emphasized by the Standards for Educational and Psychological Testing, requires evidence supporting interpretations made from test scores (American Educational Research Association et al., 2014). This principle, in AI-enabled environments, requires universities to examine whether assessment tasks still measure intended constructs or whether they inadvertently measure students' ability to access and manipulate AI-generated responses.

Technology has been shown by research on AI-supported assessment to be capable of enhancing assessment efficiency through automated scoring, adaptive testing, learning analytics, and personalized feedback (Holmes et al., 2022; OECD, 2021). Scholars, however, caution that efficiency gains must not replace human judgement, professional oversight, and careful validation of assessment processes.

A dual reality is therefore revealed by the literature: AI has the potential to strengthen assessment systems, but poorly governed AI integration may weaken assessment credibility. The importance of combining technological innovation with strong principles of educational measurement is reinforced by this tension.

6.3 African and Nigerian Perspectives on Generative Artificial Intelligence in Higher Education

Compared with scholarship from Europe, North America, and parts of Asia, research on GenAI in African higher education remains relatively emerging. Available studies, however, suggest that African universities face distinctive opportunities and challenges in adopting AI technologies. UNESCO(2023) notes that developing educational systems require context-sensitive approaches to AI integration because technological adoption is influenced by infrastructure, policy environments, digital inequalities, and institutional capacity. These factors are particularly relevant to African higher education systems where disparities in technological access may influence both opportunities for AI-supported learning and risks of educational inequality.

Within Nigeria, discussions surrounding AI in higher education have increased following the widespread availability of ChatGPT and related tools. Existing scholarship indicates growing awareness among lecturers and students regarding the educational possibilities of GenAI, including academic writing support, research assistance, and learning enhancement. However, concerns remain regarding academic dishonesty, insufficient AI policies, limited digital infrastructure, and inadequate AI literacy.

The Nigerian university system presents a particularly important context because institutions are expected to participate in global digital transformation while addressing longstanding challenges related to educational quality assurance, examination integrity, and unequal technological resources. Assessment responses to GenAI must therefore consider Nigerian realities rather than simply adopting approaches developed in technologically advanced contexts.

6.4 Research Gap

Several important gaps requiring further scholarly attention have revealed by the literature reviewed.

First, though it is true that international research on GenAI and higher education is expanding rapidly, much of the available evidence however originates from developed educational systems. There is need for additional context-specific scholarship because the technological conditions, institutional resources, and policy environments of these contexts differ significantly from those of Nigerian universities.

Second, attention in existing research has lately been focused on AI adoption, academic integrity concerns, and general educational implications. Examining GenAI through the specific theoretical lens of Educational Measurement and Evaluation has comparatively received less attention, particularly the variables of assessment validity, authenticity, reliability, fairness, and transparency.

Third, assessment integrity is treated primarily in many existing discussions as a problem of academic misconduct. The emergence of GenAI, however, necessitates a broader conceptualization of integrity that includes the capacity of assessment systems to generate valid and defensible evidence of learning.

Fourth, Nigerian higher education literature regarding institutional preparedness, AI literacy, and assessment redesign strategies is still limited. The conceptual guidance on how universities can balance responsible AI adoption with preservation of educational standards is insufficient.

These gaps are addressed in this article by synthesizing current scholarship and proposing a framework for understanding GenAI-responsive assessment integrity within Nigerian universities. The framework emphatically asserts that sustainable assessment quality requires the interaction of technological competence, ethical governance, authentic assessment design, and validity-focused evaluation.

7. DISCUSSION

A significant turning point has been recorded in the history of educational assessment with the emergence of Generative Artificial Intelligence (GenAI). From the evidence reviewed in this article, it is certain that GenAI is neither merely a threat to academic integrity nor simply a technological solution to educational challenges. It rather represents a transformative force requiring universities to reconsider how learning is demonstrated, measured, interpreted, and certified.

The conceptual framework developed in this article clarifies that the relationship between GenAI and assessment integrity is neither direct nor deterministic. GenAI adoption is influenced by perceived usefulness and perceived ease of use, but the educational consequences of adoption depend on AI literacy, institutional preparedness, governance and responsible utilization. These conditions subsequently create a need for assessment transformation, while the validity of the transformed assessment system must remain subject to continuous measurement scrutiny. Assessment integrity therefore emerges from the interaction of technological, institutional, pedagogical and measurement conditions rather than from the presence or absence of AI alone.

The most important challenge posed by GenAI is not the existence of artificial intelligence itself but the capacity of assessment systems to continue producing valid and trustworthy evidence of student learning. That is the central argument emerging from this review, a position that aligns with Messick's (1989) unified theory of validity, which emphasizes that the meaning and consequences of assessment scores must be carefully justified. Assessment practices in AI-enabled environments must therefore be examined continuously to ensure that scores represent students' actual knowledge, skills, and competencies rather than their ability to obtain unacknowledged technological assistance.

The literature reviewed attest to the fact that traditional assessment approaches are increasingly vulnerable in environments where AI systems can generate sophisticated academic responses. Written assignments, essays, and take-home examinations must still be acknowledged as remaining valuable educational tools; nevertheless, when used without modification, they may provide limited evidence of individual student competence. This challenge does not imply that such assessments should be abandoned; it rather implies that they should be redesigned to incorporate processes that reveal students' reasoning, reflection, and ownership of learning.

The growing scholarly emphasis on authentic assessment is reinforced by this finding. Application knowledge, solving complex problems, justifying decisions, and demonstrating competence within meaningful contexts are some of the basic requirements of authentic assessment approaches. Such approaches, because they focus less on the production of information, and more on the demonstration of understanding, creativity, critical judgement, and professional capability, are particularly suitable for AI-enabled environments. Why stakeholders' perceptions of GenAI influence its educational impact is further explained by the integration of the Technology Acceptance Model(Davis, 1989). Where lecturers and students perceive AI technologies as useful and accessible, adoption is likely to increase. Technological acceptance alone, however, does not guarantee educational quality. Institutional policies, AI literacy, ethical awareness, and assessment practices capable of preserving validity, are all required for responsible adoption.

The importance of institutional preparedness is also highlighted in the review. GenAI cannot be managed by universities through prohibition alone. Given that AI technologies are becoming embedded within academic activities, institutions require comprehensive governance frameworks that define appropriate use, promote transparency, protect assessment fairness, and support responsible innovation.

Institutional preparedness is particularly important for Nigerian universities. While opportunities for improving educational access, research productivity, and learning support are offered by GenAI, unequal technological infrastructure may create new forms of educational inequality. It is therefore imperative that universities consider issues of digital inclusion, staff capacity development, student support, and sustainable technology policies.

It is also revealed in the discussion that AI literacy has become an essential component of contemporary educational competence. For lecturers, the ability to redesign assessments, evaluate AI-assisted work, and guide students towards ethical AI practices is required. It is also exigent for students to acquire knowledge of responsible AI use, including transparency, critical evaluation of AI-generated information, and recognize the limits of automated systems.

The conceptual contribution of this article can be found in the proposition of GenAI-responsive assessment integrity. This concept does not imply or suggest the replacement of established theories of assessment but rather, the extension of current understandings of integrity to accommodate technological transformation. The concept views assessment integrity as a dynamic quality maintained through the interaction of valid assessment design, responsible AI adoption, institutional governance, and stakeholder competence.

8. IMPLICATIONS FOR EDUCATIONAL MEASUREMENT AND EVALUATION

Several important implications for the discipline of Educational Measurement and Evaluation are entailed by the emergence of GenAI.

To start with, measurement specialists must reconsider traditional assumptions regarding assessment design. Greater emphasis on assessments that capture complex human abilities, including reasoning, creativity, judgement, problem-solving, and application of knowledge is required by the increasing availability of AI-generated content.

Second, assessment validity must remain at the centre of educational decision-making. The need for measurement professionals to examine whether assessment tasks continue to represent intended constructs and identify possible sources of construct-irrelevant variance introduced by inappropriate AI use, is underscored.

Renewed attention is required for reliability and fairness considerations. Assessment systems should guard against the evaluation of students under unequal conditions, ensure that students are evaluated under equitable conditions and ensure that differences in access to AI technologies do not unfairly advantage or disadvantage particular groups.

Fourth, transparency as a principle should become fundamental in AI-enabled assessment. Clear expectations

regarding when and how students may use GenAI tools and how such use should be acknowledged should be established by universities.

Finally, future researchers, teachers, and assessment specialists should be prepared in Education Measurement and Evaluation programmes, to engage with emerging technologies.

Professional preparation should increasingly incorporate AI literacy, digital assessment, ethical technology integration, and authentic assessment design as its important components.

9. CONCLUSION

Generative Artificial Intelligence has fundamentally altered the educational landscape by challenging established approaches to knowledge production, learning, and assessment. This article has demonstrated that the primary challenge facing universities is not simply preventing AI use but ensuring that assessment systems continue to generate credible evidence of student learning.

The review established that assessment integrity in the era of GenAI requires a broader understanding that incorporates validity, authenticity, reliability, fairness, transparency, and defensible interpretation of assessment outcomes. While AI technologies create legitimate concerns regarding academic integrity, they also provide opportunities for assessment innovation when integrated responsibly.

The future of educational assessment depends on a balanced approach that combines technological acceptance with measurement principles. Universities must therefore move beyond detection-focused strategies towards assessment redesign, AI literacy development, institutional governance, and human-centred approaches to educational innovation.

For Nigerian universities, this transition requires context-sensitive policies, investment in digital capacity, professional development for lecturers, and continuous improvement of assessment practices. By adopting responsible approaches to GenAI integration, universities can preserve the credibility of academic qualifications while preparing learners for participation in an increasingly AI-driven society.

Ultimately, GenAI should be understood not as an enemy of educational assessment but as a catalyst for improving how learning is measured and valued. The institutions that successfully integrate technological innovation with sound principles of Educational Measurement and Evaluation will be best positioned to maintain educational quality and public confidence in higher education.

10. RECOMMENDATIONS

The following recommendations are proposed, based on the review:

Comprehensive institutional policies regulating the ethical and responsible use of GenAI in teaching, learning, research, and assessment, should be developed by Nigerian universities.

To incorporate authentic, performance-based, and process-oriented approaches capable of evaluating higher-order competencies, assessment systems should be redesigned.

Continuous AI literacy programmes for lecturers, students, and administrators should be established by universities.

Active roles in developing AI-responsive assessment frameworks should be played by Educational Measurement and Evaluation specialists.

National guidance on responsible AI integration in higher education should be provided by the National Universities Commission and relevant educational authorities.

Strengthening digital infrastructure and promoting equitable access to AI-supported learning resources should be guaranteed by universities.

Transparency regarding the use of AI tools in academic work should be required in assessment policies.

Empirical relationships among AI literacy, institutional preparedness, assessment practices, and student learning outcomes within Nigerian universities should be investigated by further research.

REFERENCES

1. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.
2. Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52–62. https://doi.org/10.61969/jai.1337500
3. Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. SAGE Publications.
4. Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating? Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148
5. Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.
6. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008
7. Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Addison-Wesley.
8. Holmes, W., Bialik, M., & Fadel, C. (2022). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
9. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274
10. Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13–103). American Council on Education and Macmillan.
11. OECD. (2021). Digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots. OECD Publishing. https://doi.org/10.1787/589b283f-en
12. Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching & Learning Practice, 20(2), Article 07. https://doi.org/10.53761/1.20.02.07
13. UNESCO. (2021). AI and education: Guidance for policy-makers. UNESCO.
14. UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
15. Xia, Q., Chiu, T. K. F., Zhou, X., & Chai, C. S. (2024). Systematic review of research on artificial intelligence in higher education assessment. International Journal of Educational Technology in Higher Education, 21, Article 34.
16. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0