



THEMATIC ANALYSIS OF POLICY DOCUMENTS AND EXAMINATION INSTRUMENTS ON TESTS AND MEASUREMENT IN NIGERIAN EDUCATION

Oyeronke Christiana Paramole¹, Moses Adeleke Adeoye²

^{1,2}Al-Hikmah University, Nigeria

Email: princeadelekm@gmail.com

ABSTRACT

Testing and measurement play a central role in determining learning outcomes and guiding instructional decision-making within educational systems. In Nigeria, national policy documents emphasise continuous, holistic, and development-oriented assessment, yet concerns persist regarding the extent to which these policy intentions are reflected in examination practices. This study examines how tests and measurement are conceptualised in key Nigerian educational policy documents and determines the degree of alignment between these policy expectations and the examination instruments used by major national examination bodies. A qualitative research design was employed, using thematic analysis to review selected policy documents and examination instruments from WAEC, NECO, and JAMB. The analysis revealed that although policy frameworks promote assessment across cognitive, affective, and psychomotor domains, examination practices remain predominantly summative and cognitively focused, with limited evidence of holistic or continuous assessment approaches. The findings also highlighted issues relating to validity, fairness, and teacher capacity in the implementation of assessment requirements. The study concludes that there is a clear gap between policy intentions and actual assessment practice in Nigeria's education system. The research contributes to ongoing discourse on assessment reform by demonstrating the need for strengthened teacher assessment literacy, clearer operational guidelines for continuous assessment, and improved quality assurance processes in examination systems.

Keywords: Tests and Measurement, Continuous Assessment, Educational Policy, Examination Practices, Teacher Assessment Literacy, Validity and Fairness, Nigerian Education System

*Corresponding Author: princeadelekm@gmail.com

Received: November 7th 2025; Revised: June 4th 2026; Accepted: June 30th 2026

DOI: <https://doi.org/10.34125/jerit.v3i1.44>

Reference to this paper should be made as follows: Paramole, O. C, & Adeoye, M. A. A Thematic Analysis of Policy Documents and Examination Instruments on Tests and Measurement in Nigerian Education. *JERIT: Journal of Educational Research and Innovation Technology*, 3 (1), 16-34.

E-ISSN: [3063-5462](https://doi.org/10.3063-5462)

Published by: JERIT: Journal of Educational Research and Innovation Technology

INTRODUCTION

In contemporary educational systems, the processes of testing, measurement, assessment and evaluation serve as crucial mechanisms for monitoring learning, shaping instruction, and informing policy decisions (Kubiszyn & Borich, 2024). In the Nigerian system, these processes have distinct historical and contextual dimensions,

shaped by national education policies, examination practices, and teacher competencies. It is therefore imperative to situate the present study within that broader landscape of “tests and measurement” in Nigeria. Firstly, tests are often defined as instruments or tasks that elicit responses from learners to infer levels of knowledge, skills or attitudes (Adom et al., 2020). Measurement is the assignment of numbers or scores to those responses, thereby transforming qualitative behavioural data into quantitative form (Dalati, 2018). Assessment then refers to the process of collecting and interpreting those scores, and evaluation is the broader judgment about the educational value or significance of those scores (Mohan, 2023). Within Nigeria, this chain forms the backbone of what many policy documents and teacher-training programmes call “continuous assessment” and external examinations. The role of external examinations and external measurement instruments in Nigeria has been especially significant. For example, high-stakes tests such as those administered by the West African Examinations Council (WAEC), National Examinations Council (NECO), and Joint Admissions and Matriculation Board (JAMB) form major gateways into further education, selection into universities, and certification of learning (Balogun, 2023; Raji & Baidoo-Anu, 2025). These assessments thereby shape not only student outcomes but also institutional behaviours, curriculum orientations, and teaching practices.

Nevertheless, despite the centrality of measurement and testing, scholars have documented persistent challenges in the Nigerian context. For instance, Katsayal et al. (2024) found that many classroom-based achievement tests lacked adequate psychometric properties, indicating a deficiency in teacher competency in test construction. Further, Adeosun (2019) revealed that many teachers hold negative attitudes towards validation of non-standardised achievement tests, which undermines the legitimacy of measurement practices in secondary schools. At the level of high-stakes examinations, Metibemu et al. (2024) revealed on an NECO Physics paper found evidence of differential item functioning (DIF) across northern and southern candidate groups, suggesting the test instruments may not equally serve all examinees. These findings indicate that measurement practices in Nigeria are confronted by issues of validity, fairness, reliability, and alignment – all key constructs in measurement theory. Moreover, the teacher preparation and policy environment also warrant attention. Literature indicates that teacher education in Nigeria includes courses on educational measurement and test construction but the implementation of such measurement skills in practice appears weak (Atakpo, 2017; Ibrahim, 2020; Ogunyinka et al., 2015). There is therefore a disjunction between the ideal forms of measurement and the praxis of assessment within Nigerian classrooms.

From a policy perspective, Nigeria’s educational regulatory frameworks have repeatedly emphasised continuous assessment, quality assurance, and standardised external examinations as part of the drive to improve educational outcomes (Ezewuzie et al., 2025; Hassan, 2025; Okeke & Chukwudebelu, 2024). Yet, the translation of these intentions into the construction, administration and measurement of tests has been uneven. Given this context, it becomes evident why a study of the thematic structure of policy documents and examination instruments on tests and measurement is timely and necessary. By focusing on how policy texts articulate tests and measurement, and

how examination instruments instantiate or diverge from those articulations, we can surface the patterns of meaning, regulation and practice that shape assessment culture in Nigeria. These points set the stage for investigating how tests and measurement are constructed, presented and regulated within the Nigerian education system – via an analysis of policy documents and examination instruments.

Problem Statement

Despite the lofty intentions laid out in Nigeria’s education policy documents – such as the insistence on fair, valid, reliable, and equitable assessment systems – there remains a persistent and troubling discrepancy between what is promised and what is practised in the realm of tests and measurement. This gulf raises significant concerns about the effectiveness of assessment as a driver of educational quality and equity in the Nigerian context. Policy frameworks in Nigeria articulate an ambitious vision for assessment. For example, the mantra of continuous assessment was introduced to move beyond one-off high-stakes examinations and to provide a more holistic, ongoing appraisal of learner performance (Raji, 2023). This policy ideal emphasises that assessments should capture cognitive, affective and psychomotor domains, facilitate feedback, promote learning and reduce over-dependence on terminal exams (Mohammed et al., 2024). Moreover, the notion that tests should be constructed with attention to measurement principles – validity, reliability, fairness, standardisation – appears in broader educational measurement literature, which argues that sound test construction supports valid inferences about students’ achievement and learning (Omachi, 2019).

Yet, in practice, multiple studies have documented shortcomings in how tests and measurements are implemented in Nigerian schools and examination systems. Adeosun (2019) found that secondary-school teachers held predominantly negative attitudes toward the validation of non-standardised achievement tests, signalling weak fidelity to measurement standards in classrooms. Another investigation into differential item functioning in a Nigerian national examinations body revealed biases in item performance across groups of examinees – undermining the fairness and validity of major examinations (Adediwura & Asowo, 2021). Furthermore, continuous assessment practices sometimes remain rhetorical: although policy expects periodic, comprehensive assessment of learning across domains, research indicates that implementation is inconsistent and heavily reliant on conventional terminal tests (Raji, 2023). The juxtaposition of policy-driven ideals and the empirical realities suggests a clear gap. On one side, the formal regulatory documents promise assessment systems anchored in measurement theory and progressive educational values. On the other side: actual exam instruments and classroom practices reveal patterns of misalignment – tests that may lack standardisation, measurement rigour, domain coverage, or fairness. This gap yields several problematic implications: Validity threat: If examination items do not reflect the intended constructs (e.g., critical thinking, application, affective and psychomotor domains), then the inferences drawn from test scores are questionable. Reliability concerns: Without consistent application of measurement standards, scores may vary not because of differences in learner ability but because of variability in test construction and administration. Equity issues:

Biases in test items undermine fairness—students from different regions or backgrounds may be unfairly advantaged or disadvantaged. Policy-practice disjuncture: When what is policy is not mirrored in practice, the credibility of the measurement system and the accountability function of assessment are weakened.

Given this scenario, there is a compelling need to systematically examine how policy documents articulate tests and measurement, and how these articulations compare with the actual instruments used in examination bodies. By doing so, the study sheds light on the meaning-making of measurement constructs in both the regulatory texts and the examination practices—revealing where alignment exists, where divergence occurs, and why such divergence matters. This disparity between intention and implementation constitutes the central problem guiding this study, and warrants a thematic exploration of policy texts and examination instruments to uncover patterns of alignment, divergence and implication.

METHODS

This study employed a qualitative research design anchored in thematic analysis. The choice of this design was motivated by the need to interpret, rather than quantify, how tests and measurement are conceptualised, regulated, and practised within the Nigerian education system. Rather than seeking statistical generalisations, the focus was on uncovering patterns of meaning embedded in official policies and examination instruments. Thematic analysis provided a systematic and flexible way to identify, organise, and interpret recurring themes across multiple texts, enabling a critical comparison between policy ideals and actual assessment practices. Two categories of documentary evidence formed the primary data set: Policy Documents: These included official texts issued by government ministries, teacher education regulators, and examination bodies. Specifically, the following were consulted: The National Policy on Education (latest edition, Federal Ministry of Education); Continuous Assessment Implementation Guidelines (Federal and selected State Ministries of Education); Teacher Education Curricula (National Commission for Colleges of Education and selected universities' Faculties of Education) and WAEC, NECO, and JAMB regulations, syllabi, and test specifications. These documents provided insight into how educational authorities formally conceptualise tests, assessment standards, and principles of measurement. Past question papers and marking guides from WAEC, NECO, and JAMB covering English, Mathematics, and selected science subjects between 2013 and 2025. These instruments reflected the practical implementation of testing principles and offered a lens for comparing official expectations with examination realities. A purposive sampling strategy was adopted to ensure relevance and depth of analysis. The inclusion criteria were: Documents issued by government ministries, teacher education regulators, or recognised examination bodies. Coverage of dominant subjects (English, Mathematics, Sciences) and major examination types central to Nigerian education. Texts and examination papers dated between 2013 and 2024 to capture current practices and align with ongoing reforms. Based on these criteria, a total of 12 policy documents and 18 examination instruments were selected for analysis. Policy documents were obtained through multiple channels, including the official websites of WAEC, NECO, JAMB, and the Federal Ministry of Education. Additional materials such as teacher training

curricula were sourced from university libraries and faculties of education. Examination instruments were accessed from examination board archives and selected secondary schools. All documents were converted into digital formats (PDF or Word) to facilitate systematic analysis. The data were analysed using the six-step framework for thematic analysis outlined (Kiger & Varpio, 2020). This involved: Familiarisation: Repeated reading of the documents to gain an initial sense of the content. Generating Initial Codes: Segments of text related to tests, measurement standards, assessment policies, and practices were coded. Searching for Themes: Codes were grouped into broader categories reflecting key concerns (e.g., fairness, validity, continuous assessment, examination standards). Reviewing Themes: Themes were checked against the data to ensure consistency and clarity. Defining and Naming Themes: Overarching ideas and sub-themes were identified and labelled to capture their essence. Producing the Report: Themes were organised into findings that highlighted convergences and divergences between policy ideals and examination practices. NVivo 14 software was used to manage, code, and organise the data. The software facilitated systematic comparison across multiple sources, ensuring transparency and rigour in the analysis. Since the study relied exclusively on publicly available and institutional documents, no direct ethical risks were involved. However, due diligence was observed in citing and referencing all sources appropriately. Sensitive data, such as student scripts or restricted internal reports, were deliberately excluded to maintain professional and ethical standards.

RESULTS AND DISCUSSION

Research Finding One: Themes from Policy Documents

The analysis of Nigerian educational policy documents revealed several recurring and interrelated themes that reflect the state's formal conceptualisation of tests, measurement, and assessment. These themes highlight the normative ideals that underpin the design, regulation, and implementation of educational assessment practices in Nigeria. Three dominant thematic categories emerged: The conceptualisation of testing and measurement, The policy vision of continuous assessment, and The standards of fairness, validity, and reliability as mandated across national frameworks. Each of these themes illustrates how educational authorities define and regulate testing practices within Nigeria's broader goals of quality assurance, accountability, and learner development.

Conceptualisation of Testing and Measurement

Policy documents such as the National Policy on Education conceptualise testing and measurement primarily as mechanisms for monitoring learning outcomes and ensuring national educational standards. The policy defines educational assessment as a process for "measuring learners' progress in cognitive, affective, and psychomotor domains," thereby positioning measurement as a holistic and diagnostic tool rather than a purely summative one. This conception reflects an integrated measurement philosophy influenced by international best practices, but also grounded in the local goal of maintaining uniform standards across Nigeria's diverse educational regions. The Federal Ministry of Education (FME) and the National Commission for Colleges of Education (NCCE) emphasise that measurement is not

only a technical activity but also an ethical responsibility – to produce results that are valid, fair, and reflective of true learning outcomes.

However, the policy language often blurs distinctions between testing, measurement, and evaluation, using the terms interchangeably. This terminological ambiguity, though common in developing education systems, may contribute to misinterpretation during implementation. While testing is defined operationally (e.g., as examination procedures), measurement is discussed philosophically (as quantification of learning), and evaluation encompasses judgement and policy decisions. The documents reveal an aspirational but under-defined conceptual framework—one that acknowledges the importance of psychometric integrity but provides limited operational guidance on achieving it. The Teacher Education Curriculum reinforces this conceptualisation by mandating that pre-service teachers “demonstrate competence in test construction, administration, and interpretation of results.” This reinforces the state’s recognition of teacher assessment literacy as a foundational element in improving the quality of testing and measurement. Yet, the curriculum provides minimal procedural clarity on how these competencies are to be evaluated in practice—leaving room for local variations in interpretation and application.

The Policy Vision of Continuous Assessment Ideals

Across policy texts—particularly the Continuous Assessment Implementation Guidelines and state-level directives—a strong rhetorical commitment to continuous assessment (CA) was evident. Continuous assessment is described as a comprehensive and cumulative process that evaluates “all aspects of a learner’s growth over time,” including cognitive, affective, and psychomotor skills. It is positioned as a corrective to the limitations of terminal examinations and a means of ensuring fairness, inclusivity, and ongoing feedback in learning. The ideal of continuous assessment within policy discourse is built upon several interlocking principles: Assessment should extend beyond cognitive achievement to include affective and psychomotor development. Learner progress should be tracked progressively over time rather than captured in a single examination event. Assessment outcomes should guide instruction and support individual learner improvement. CA is intended to reduce the disadvantage of one-time high-stakes testing, thus promoting fairness across learners. Nevertheless, a close reading of these documents reveals implementation tensions. While the guidelines articulate detailed expectations, they are less explicit about monitoring mechanisms, teacher capacity development, and assessment moderation processes. The policy language assumes a level of infrastructural readiness and teacher competence that, as subsequent empirical studies show (Adedoyin et al., 2024; Obiakor, 2024), may not be uniformly present across Nigerian schools. Moreover, policy provisions treat continuous assessment as complementary to, rather than replacing, summative examinations—suggesting a dual-assessment model. This hybridisation reflects an effort to balance local traditions of examination-driven accountability with global trends toward formative assessment. Yet, without robust coordination between classroom-based CA and external examinations, policy coherence remains limited.

Standards of Fairness, Validity, and Reliability

A recurrent and normative theme across the policy corpus is the emphasis on the principles of fairness, validity, reliability, and standardisation in all forms of educational assessment. The WAEC and NECO Examination Regulations, for instance, explicitly reference the need for test items to meet established psychometric standards. Fairness is defined in terms of equitable access, linguistic clarity, and the absence of cultural bias in examination items. Validity is described as the extent to which test results accurately reflect intended learning outcomes, while reliability refers to consistency across administrations and marking processes. Similarly, JAMB's Unified Tertiary Matriculation Examination (UTME) Guidelines highlight validity and reliability as core principles for maintaining national comparability in selection procedures. However, while these documents affirm these standards rhetorically, they often lack transparency about how validity evidence is gathered or how reliability checks are operationalised. From a thematic standpoint, policy emphasis on fairness and validity demonstrates a clear aspirational standardisation agenda—an attempt to align Nigeria's assessment system with international best practices. Yet, the absence of detailed frameworks for item review, moderation, and validation suggests that these standards function more as policy ideals than as systemic guarantees. Furthermore, policy references to fairness tend to focus on procedural equality (e.g., access to examination opportunities) rather than on substantive fairness—that is, addressing the socio-economic, linguistic, or infrastructural inequities that affect learners' performance. This narrower conceptualisation limits the policy's capacity to address deeper structural dimensions of fairness in educational measurement.

These findings highlight the gap between policy articulation and practice enactment, setting the stage for the next phase of analysis: a comparison of these policy ideals with the empirical realities reflected in actual examination instruments (WAEC, NECO, and JAMB papers and marking guides). The juxtaposition will illuminate how far the principles of validity, reliability, and fairness are realised—or compromised—in practice.

Research Finding Two: Themes from Examination Instruments

The second phase of the analysis involved an in-depth thematic examination of selected examination instruments, including test papers, marking schemes, and examination guidelines issued by national and subnational examining bodies such as the West African Examinations Council (WAEC), the National Examinations Council (NECO), and the Joint Admissions and Matriculation Board (JAMB). Additionally, sample internal school-based tests and continuous assessment records were reviewed to capture how policy ideals are translated into practical assessment instruments within the Nigerian educational system. The analysis revealed three dominant themes reflecting the relationship between policy intentions and actual examination practices: Test design patterns and the dominance of cognitive measurement, Partial alignment with policy standards on validity and fairness, and Deviation from continuous assessment ideals and holistic evaluation. Each theme highlights the gap between policy prescriptions and the realities of examination implementation, providing insight into how the ideals of fairness, validity, and comprehensiveness are enacted—or compromised—in practice.

Test Design Patterns and the Dominance of Cognitive Measurement

A major theme emerging from the analysis is the overwhelming emphasis on cognitive learning outcomes in the design of examination instruments. Across all examined samples—particularly WAEC and NECO question papers—test items predominantly assess recall, comprehension, and application-level skills, aligning primarily with the lower tiers of Bloom’s Taxonomy. Items measuring affective or psychomotor domains were almost absent. For example, in Senior School Certificate Examination (SSCE) papers, the majority of questions required factual recall, short explanations, or structured essay responses, while very few demanded critical analysis, synthesis, or problem-solving in authentic contexts. This design trend suggests a narrow operationalisation of “educational achievement,” focusing more on content mastery than on competence development, a concern similarly reported by JAMB (2023) in their critique of assessment trends in Nigerian schools. Although policy documents emphasise that assessment should be holistic—covering cognitive, affective, and psychomotor domains—the test designs analysed show a systemic bias towards written, summative, and cognitive-heavy testing formats. This dominance of knowledge-based evaluation limits the diagnostic and formative potential of assessments, reinforcing the “examination-oriented” culture that scholars like Obilor and Akwawo (2024) and Kuku and Alade (2017) describe as a structural weakness in Nigeria’s assessment system. The marking schemes accompanying these tests also reveal a rigid, point-based allocation system that rewards factual precision over interpretive depth or creative reasoning. Consequently, while objectivity and reliability are maintained, the validity of these instruments—as measures of comprehensive educational attainment—remains questionable.

Partial Alignment with Policy Standards on Validity and Fairness

The second prominent theme concerns the partial alignment of examination instruments with policy standards related to fairness, validity, and reliability. Examination bodies such as WAEC and JAMB explicitly reference psychometric quality standards in their technical manuals and public statements, aligning in principle with policy expectations. However, closer analysis of test items, linguistic presentation, and content distribution reveals uneven adherence to these standards. In sciences and mathematics, question construction follows established syllabi and standardisation principles, with items typically undergoing pretesting, moderation, and review to ensure content validity and reliability, as noted in WAEC Chief Examiners’ Reports (Oladele, 2021). However, in other subjects—especially humanities and social sciences—questions frequently demonstrate content overlap, redundant testing of the same skill, and limited representation of higher-order thinking, indicating inconsistencies in operational validity.

Furthermore, fairness remains a contested issue. While the language of test items is generally formal and uniform, linguistic complexity and regional references occasionally disadvantage candidates from rural or linguistically diverse backgrounds. The absence of contextual adaptation or differentiated item design suggests a limited conceptualisation of fairness, focusing mainly on procedural uniformity rather than equitable accessibility. Another observable trend is the absence of transparent feedback mechanisms. Unlike formative assessments envisioned in

policy, external examinations provide no structured feedback to learners or schools. Results are summarised numerically, with little diagnostic information on strengths or weaknesses. This practice, while maintaining administrative efficiency, contradicts the feedback and improvement ideals embedded in Nigeria's Continuous Assessment Guidelines. Examination bodies ostensibly adhere to psychometric standards; however, their practices demonstrate a partial and inconsistent application of these standards, illustrating the "policy-practice gap" identified by Raji and Baidoo-Anu (2025) in institutionalised assessment systems.

Deviation from Continuous Assessment Ideals and Holistic Evaluation

The third theme identifies a systematic deviation from the continuous assessment (CA) ideals articulated in Nigeria's education policy. Policy frameworks envision CA as a process for monitoring a learner's growth across the cognitive, affective, and psychomotor domains throughout the school year. However, the examination instruments analysed reflect minimal integration of CA principles. School-based assessments often serve as miniature replications of public examinations, focusing on recall-oriented written tests rather than diversified modes such as portfolios, projects, observations, and oral assessments. Teachers' CA records – when available – show a limited variety in assessment types, dominated by periodic tests rather than continuous observation or formative evaluation. This pattern echoes findings by Bolaji et al. (2017), who observed that CA in Nigerian schools is often reduced to a bureaucratic routine rather than a formative learning process. Consequently, the intended synergy between classroom-based assessment and external examination systems remains largely theoretical. The disconnection between these two levels of evaluation undermines the validity of the overall assessment framework and perpetuates what Raji (2023) calls "the myth of measurement objectivity without pedagogical depth."

Furthermore, the psychometric orientation of national examinations reinforces a culture of summative finality rather than formative continuity. Examinations are treated as end-points for selection and certification, not as integral components of the teaching-learning process. The thematic evidence thus demonstrates a misalignment between policy vision and institutionalised assessment culture, where measurement serves administrative rather than pedagogical ends. The thematic analysis of examination instruments thus exposes a structural misalignment between the policy ideals of comprehensive, fair, and developmental assessment, and the actual design and implementation of examinations in Nigeria. Three interpretive conclusions emerge: Examination instruments remain predominantly cognitive and summative, reflecting a narrow interpretation of educational achievement. Psychometric standards of validity and fairness are only partially applied, leading to inconsistencies across subjects and contexts. Continuous assessment ideals have not been effectively operationalised, resulting in a fragmented assessment system disconnected from everyday learning processes. Overall, while Nigerian assessment policy frameworks advocate modern, student-centred, and continuous evaluation practices, examination instruments continue to reflect traditional, test-oriented, and gatekeeping functions. This finding underscores the persistent policy-practice gap in Nigerian education,

where progressive assessment philosophy coexists with conservative testing traditions.

Research Finding Three: Comparative Insights between Policy Expectations and Practice Realities

The third phase of the analysis entailed a cross-thematic comparison of data emerging from policy documents and examination instruments. This stage sought to uncover how the conceptual ideals of testing, measurement, and assessment articulated in Nigerian education policies are reflected—or contradicted—in actual examination practices. The analysis revealed both convergences and divergences between policy intentions and practical realities. While national policies espouse principles of validity, fairness, continuous assessment, and holistic evaluation, the evidence from examination instruments suggests a partial translation of these ideals, constrained by institutional culture, examination pressures, and systemic inertia. Three major comparative themes emerged from the analysis: Convergence in procedural standardisation but divergence in assessment philosophy, and Alignment in policy rhetoric but contradiction in implementation depth.

Convergence in Procedural Standardisation but Divergence in Assessment Philosophy

Both policy documents and examination instruments display a shared commitment to procedural standardisation—that is, ensuring fairness, reliability, and uniformity in assessment processes. This convergence is visible in the explicit mention of psychometric principles in policy frameworks such as the National Policy on Education, Continuous Assessment Implementation Guidelines, and WAEC/NECO regulatory manuals. Similarly, examination bodies demonstrate awareness of these procedural norms through practices such as moderation, test blueprints, and uniformity in marking schemes. This alignment indicates institutional consciousness of the technical demands of valid measurement, which resonates with international best practices in assessment design. However, the philosophical divergence between policy and practice lies in the underlying purpose of assessment. Policy documents frame testing as a tool for learning, diagnosis, and improvement, grounded in a formative educational philosophy. In contrast, examination instruments operationalise assessment as a mechanism for selection, certification, and exclusion. The analysis revealed that while procedures are standardised, the pedagogical intent of assessment has been lost in translation. As Onyeabo (2021) notes in his “policy enactment” theory, the meaning of policy often shifts as it moves from textual formulation to institutional implementation—a phenomenon evident in how Nigerian testing practices have institutionalised procedural compliance without embodying policy-driven learning goals.

Alignment in Policy Rhetoric but Contradiction in Implementation Depth

Another key comparative insight is the strong alignment in policy rhetoric with international standards of fairness, inclusivity, and continuous assessment—yet a contradiction in actual implementation depth and consistency. The National Policy on Education and teacher education curricula articulate clear expectations for multi-dimensional evaluation across cognitive, affective, and psychomotor domains. These ideals align with global frameworks on formative assessment by Black assessment for

learning (Black, 2015). However, in practice, the examination instruments reviewed remain overwhelmingly restricted to cognitive recall, with minimal evidence of domain integration. This contradiction suggests a symbolic rather than functional policy alignment—what Moore and Paris (2019) describe as “rhetorical coherence without operational coherence.” While policymakers adopt the language of innovation, the structural practices of examination bodies remain anchored in traditional paper-and-pencil formats, high-stakes selection processes, and rigid scoring systems. Furthermore, whereas policy texts advocate learner-centred and diagnostic approaches, the test instruments reveal a teacher- and examiner-centred orientation, focused on grading rather than learning progression. This disparity reflects what Jenkins and Gray (2020) term “policy idealism confronting assessment realism”—a condition where progressive educational visions struggle against entrenched bureaucratic and logistical constraints.

Research Finding Four: Informing Reforms in Educational Assessment and Teacher Preparation

The comparative thematic analysis of policy documents and examination instruments exposed multiple gaps between policy ideals and practical realities—particularly in how concepts such as validity, fairness, and continuous assessment are operationalised within Nigeria’s educational assessment system. The thematic patterns suggest that while Nigerian education policies are conceptually sound, their translation into practice is constrained by institutional, pedagogical, and capacity-related deficiencies. The analysis revealed that Nigeria’s assessment policies—though well-articulated—often function as symbolic frameworks rather than operational blueprints. There exists a disconnect between policy rhetoric (which promotes holistic, formative, and inclusive assessment) and the bureaucratic structures that enforce traditional high-stakes testing. This misalignment calls for evidence-based policy reform, where practice informs revision, and not merely the other way around. To achieve this, there is a need to establish periodic policy reviews grounded in empirical analysis of assessment instruments and learner outcomes. For instance, examination bodies such as WAEC, NECO, and JAMB should integrate assessment audits—structured analyses comparing examination content and practices against national standards of validity, fairness, and cognitive balance. Furthermore, integrating qualitative feedback mechanisms into the policy framework—such as post-exam reflection sessions with teachers, examiners, and students—would bridge the experiential gap between policymakers and implementers. A key finding from the research is the low assessment literacy among teachers, hindering the effective implementation of policy in classrooms. Despite courses on “Tests and Measurement” offered by the NCCE and university education faculties, their practical applications are often minimal and outdated. Teachers frequently exhibit a superficial grasp of important concepts like validity and reliability, leading to a mechanical approach to continuous assessment, focused primarily on testing rather than nurturing learning progress. This aligns with prior studies highlighting that a deeper understanding of assessment significantly impacts evaluation quality. To address this, teacher education must transition from theoretical to competency-based training, incorporating practical elements such as test construction, authentic assessment simulations, and collaborative workshops. These reforms aim to build teachers' skills

in translating policy into effective classroom practices that emphasise learning, as emphasised by Brookhart's concept of "assessment for learning."

The findings highlight the lack of sustained dialogue among policy actors, examination agencies, and teacher training institutions in Nigeria. To address this, reforms should focus on establishing collaborative platforms, such as the proposed National Assessment Reform Committee (NARC), which would include representatives from WAEC, NECO, JAMB, FME, and universities. This committee could support joint policy reviews, innovation in test design, and professional exchanges on fairness and inclusivity. By embracing a collaborative process, Nigeria can transform policy documents into adaptable frameworks that reflect classroom realities and examination experiences. Cultivating a culture of reflective and ethical assessment is essential, as current high-stakes testing practices promote a punitive view of assessment. Reforms should focus on shifting from testing to learning, emphasising fairness, inclusivity, and dignity for learners. Teacher training must incorporate reflective modules on assessment philosophy and ethics to enhance motivation, equity, and student identity. This aligns with Sustainable Development Goal 4, which advocates for inclusive and equitable quality education. Viewing assessment as a supportive process for learning can lead to more just and effective educational systems.

Discussion of findings

The study examined how tests and measurements are conceptualised within Nigerian educational policy documents and how these conceptions are reflected in the examination instruments used by major national examination bodies. The findings showed a notable gap between policy ideals and actual assessment practices. The policy documents reviewed conceptualise assessment as holistic and continuous, covering the cognitive, affective, and psychomotor domains. This aligns with the educational measurement principles that emphasise testing as a means of understanding and improving learning rather than simply ranking learners (Kubiszyn & Borich, 2024). However, while the National Policy on Education promotes continuous assessment as a core mechanism for monitoring learner development, classroom and examination practices continue to prioritise summative, end-of-term or high-stakes testing (Raji, 2023). This indicates that the intended formative purpose of assessment is not being realised in implementation. The analysis of examination instruments from WAEC, NECO, and JAMB revealed that assessment items overwhelmingly target cognitive recall and basic comprehension, with minimal emphasis on higher-order reasoning or affective and psychomotor competencies. This confirms previous findings that Nigerian assessment practices remain largely examination-driven (Balogun, 2023). Such a pattern narrows the scope of what is actually evaluated, despite policy directives requiring broader learning outcomes (Mohammed et al., 2024). The findings also show inconsistencies in validity and fairness, particularly in how examinations reflect curriculum intentions across regions. Metibemu et al. (2024) found differential item functioning in NECO Physics examinations, meaning that items did not function equivalently for learners across regions. This challenges the fairness of test outcomes and aligns with Messick's (2013) view that validity includes not only content accuracy but also the social consequences

of score use. Similarly, reliability and validity issues have been linked to teacher-constructed assessment instruments. Katsayal et al. (2024) reported that many teachers lack adequate competency in test construction, while Adeosun (2019) found negative attitudes toward validating non-standardised tests. These findings help explain why continuous assessment records often fail to reflect actual learner development. Although teacher education curricula include courses in educational measurement, implementation remains weak. Ogunyinka et al. (2015) noted that teacher development in Nigeria is hindered by insufficient practical pedagogical training. This study's findings reinforce this concern: continuous assessment in practice is frequently reduced to periodic written tests rather than a sustained, formative evaluation process (Oloba, 2023). This reflects what Ball's Policy-as-Discourse theory describes as policy recontextualisation, where frontline actors reinterpret policies under real-world constraints such as workload, accountability pressure, and lack of training. The overall pattern indicates that policy frameworks encourage comprehensive and formative assessment, but examination practices remain summative and selective. Validity and fairness concerns persist due to limited psychometric oversight and uneven teacher assessment literacy. High-stakes examinations continue to shape teaching and learning priorities, reinforcing an exam-oriented culture rather than one centred on learning development. These findings suggest that improving alignment between policy and practice will require strengthening teacher assessment literacy, enhancing moderation and validation procedures for examinations and ensuring continuous assessment is implemented as a learning support mechanism, not only as a record-keeping exercise.

CONCLUSION

This study investigated how tests and measurements are conceptualised in Nigerian education policy documents and how these policy expectations are reflected in examination instruments used by national examination bodies. The findings showed that while policy documents promote comprehensive, continuous, and developmental assessment, the actual practices within public examinations remain largely summative and focused on the cognitive domain. Examination instruments tended to emphasise recall and basic application of knowledge, with limited attention to higher-order thinking or to affective and psychomotor learning outcomes. In addition, issues relating to validity, fairness, and reliability were identified, and the capacity of teachers to implement continuous assessment in a meaningful way was found to be limited. The study concludes that there is a clear gap between assessment policy intentions and assessment practices in Nigerian education. Although the policy framework encourages holistic and continuous assessment, practice remains dominated by high-stakes examinations that prioritise testing for selection and certification. This misalignment suggests that systemic, institutional, and capacity-related factors continue to shape assessment practices more strongly than policy directives. As a result, the educational purpose of assessment—to support learning and inform instructional decision-making—is not fully realised in practice.

Recommendations

1. Policies should be supported with practical procedures and tools to guide schools and teachers in developing and using valid and reliable assessment instruments.
2. Adequate training and learning materials should be provided to support the effective implementation of assessment practices at the school and examination system levels.
3. Systematic review and refinement of examination items should be undertaken to improve validity and fairness.
4. Assessment outcomes should provide schools and learners with information that supports improvement, rather than only scores and grades.
5. Regular refresher programmes should be introduced to update teachers' knowledge and skills in classroom assessment and interpretation of assessment results.

Suggestions for Future Research

1. Studies should explore how teachers interpret and implement continuous assessment in everyday classroom contexts.
2. Comparative research across different states or regions could help identify structural factors influencing variations in assessment practice.
3. Further research should examine the impact of current assessment practices on learners' motivation, engagement, and long-term educational outcomes.

ACKNOWLEDGEMENTS

The authors acknowledge with gratitude the support received during the conduct of this research. We are particularly appreciative of the academic environment provided by Al-Hikmah University, Ilorin, which enabled access to relevant scholarly materials and institutional resources essential to the development of this study. We further acknowledge the commitment and collaborative effort shared between the co-authors throughout the research process. Each author contributed independently and equally to the conception, analysis, interpretation, and writing of this work. Lastly, we recognise the broader community of researchers whose work in the areas of educational assessment, policy analysis, and measurement provided the necessary scholarly foundation upon which this study was built.

REFERENCE

- Adediwura, A. A., & Asowo, P. (2021). Examining the Nature of Item Bias on Students' Performance in National Examinations Council (NECO) Mathematics Senior School Certificate Dichotomously Scored Items in Nigeria. *International Journal of Contemporary Education*, 5, 16. <https://doi.org/10.11114/ijce.v5i1.5402>
- Adedoyin, O. B., Altinay, F., Gemikonakli, E., Altinay, Z., & Dagli, G. (2024). National Policy on Open Educational Resources for Higher Education in Nigeria: Evaluation of Institutional Compliance Rate to Infrastructure and the Connectivity Goal. *Higher Education Policy*, 1-30. <https://doi.org/10.1057/s41307-024-00387-8>

- Adeosun, O. O. (2019). *Development and Standardisation of Early Reading Literacy Test For Pre-School Children In Oyo State, Nigeria* (Doctoral Dissertation, University of Ibadan, Nigeria).
- Adeosun, P. K. (2024). Teachers' Competency in Test Construction in Secondary Schools in Ovia North-East Local Government Area, Edo State. *Ilorin Journal of Education*, 44(1), 237-250.
- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109-119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Akanni, O. O. (2021). Evaluation of Item Response Theory in the Development and Validation of Multiple Choice Tests in Mathematics among Adolescents in Lagos Metropolis. *A Publication of Faculty of Education, Kampala International University in Tanzania*, 141.
- Alordiah, C. O. (2025). Revisiting The Concept of Reliability: A Theoretical Framework For Understanding Measurement Precision In Educational Psychology Research. *African Journal of Social and Behavioural Sciences*, 15(2).
- Alufohai, P. J., & Akinlosotu, T. N. (2016). Knowledge and Attitude of Secondary School Teachers towards Continuous Assessment Practices in Esan Central Senatorial District of Edo State. *Journal of Education and Practice*, 7(10), 71-79. <https://doi.org/10.4314/afrrrev.v10i4.10>
- Alvi, E., & Nausheen, M. (2019). Examining Grade 9 Students' Engagement in Mathematical Problem-Solving (MPS) When Working as Individuals and in Small Group Settings. *Bulletin of Education and Research*, 41(1), 163-184.
- Atakpo, T. E. (2017). *Evaluation of Early Childhood Education Teacher Preparation Programme in South East and South West Nigerian Universities* (Doctoral dissertation, University of Lagos (Nigeria)).
- Ayandele, A. A. (2017). Procedures for conducting continuous assessment in curriculum implementation for sustainable teacher education development in colleges of education in South West Nigeria. *Pacesetter: Journal of Emmanuel Alayande College of Education* 21(2), 165-177
- Balogun, J. A. (2023). *Reimagining Nigeria's Educational System: Improving Academic Performance Through High-Stakes Standardised Testing*. Routledge. <https://doi.org/10.4324/9781003388425>
- Behizadeh, N., & Neely, A. (2018). Testing injustice: Examining the consequential validity of edTPA. *Equity & Excellence in Education*, 51(3-4), 242-264. <https://doi.org/10.1080/10665684.2019.1568927>
- Black, P. (2015). Formative assessment—an optimistic but incomplete vision. *Assessment in Education: principles, policy & practice*, 22(1), 161-177. <https://doi.org/10.1080/0969594X.2014.999643>
- Bolaji, S. D., Campbell-Evans, G., & Gray, J. R. (2017). The Perils of Bureaucratic Complexity: Education for All in Nigeria. *International Journal of Educational Administration and Policy Studies*, 9(1), 1-9. <https://doi.org/10.5897/IJEAPS2016.0480>
- Bråten, I., Haverkamp, Y. E., & Anmarkrud, Ø. (2025). Gaining a deeper understanding of the deep cloze reading comprehension test: Examining

- potential contributors and consequences. *Reading and writing*, 38(2), 425-446. <https://doi.org/10.1007/s11145-024-10521-y>
- Braun, H. I., & Singer, J. D. (2019). Assessment for monitoring of education systems: International comparisons. *The ANNALS of the American Academy of Political and Social Science*, 683(1), 75-92. <https://doi.org/10.1177/0002716219843804>
- Dalati, S. (2018). Measurement and measurement scales. In *Modernising the Academic Teaching and Research Environment: Methodologies and Cases in Business Research* (pp. 79-96). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-74173-4_5
- Elnagar, A. (2021). Internationalisation policies of public schooling and neoliberal discourses: A Stephen Ball-informed analysis. *Comparative and International Education*, 50(1). <https://doi.org/10.5206/cieeci.v50i1.11075>
- Ezewuzie, S. O., Udeogu, P. A., & Okoli, H. N. (2025). Assessment of Educational Policies and Reforms In Nigeria: Challenges and Prospects. *Chukwuemeka Odumegwu Ojukwu University Journal of Arts and Social Science Education*, 4(1).
- Hassan, Z. Y. (2025). AI-Driven Horizons: Shaping the Future of Global Quality Assurance in Higher Education. *AI and Ethics, Academic Integrity and the Future of Quality Assurance in Higher Education*, 105.
- Ibrahim, M. A. (2020). Evaluation of mathematics teacher education program in a university in Kano, Nigeria. *Unpublished doctoral dissertation. University Sains Malaysia*.
- Jenkins, B., & Gray, A. (2020). Policy evaluation in British government: From idealism to realism? In *Program Evaluation and the Management of Government* (pp. 53-70). Routledge. <https://doi.org/10.4324/9780429338380-6>
- Joint Admissions and Matriculation Board (JAMB) (2023). Manual of Critical Assessment Competences for Tertiary Education Lecturers in Nigeria. Eds Oloyede, I. O., Nworgu, B. G., & Ariyo, A. O. Joint Admissions and Matriculation Board Press
- Katsayal, S. A., Mansur, U., & Shu'aibu, M. G. (2024). An Assessment of Teachers' Competency In Test Construction on Students' Performance of Junior Secondary Schools in Jigawa State. *Journal of Education Research and Library Practice* 4(8), 201-215.
- Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical teacher*, 42(8), 846-854. <https://doi.org/10.1080/0142159X.2020.1755030>
- Kubiszyn, T., & Borich, G. D. (2024). *Educational testing and measurement*. John Wiley & Sons.
- Kuku, O. O., & Alade, O. M. (2017). Impact of frequency of testing on study habits and achievement in mathematics among secondary school students in Ogun State, Nigeria. *Journal of Educational Research and Practice*, 7(1). <https://doi.org/10.5590/JERAP.2017.07.1.01>
- Levy-Feldman, I. (2025). The role of assessment in improving education and promoting educational equity. *Education Sciences*, 15(2), 224. <https://doi.org/10.3390/educsci15020224>
- Liu, L., Jones, B. F., Uzzi, B., & Wang, D. (2023). Data, measurement and empirical methods in the science of science. *Nature human behaviour*, 7(7), 1046-1058. <https://doi.org/10.1038/s41562-023-01562-4>

- Lopes, A. C. (2016). The theory of enactment by Stephen Ball: And what if the notion of discourse was different? *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 24, 1-19.
- Markus, K. A., & Borsboom, D. (2024). *Frontiers of test validity theory: Measurement, causation, and meaning*. Routledge. <https://doi.org/10.4324/9781003398219>
- Maul, A., Mari, L., Irribarra, D. T., & Wilson, M. (2018). The quality of measurement results in terms of the structural features of the measurement process. *Measurement*, 116, 611-620. <https://doi.org/10.1016/j.measurement.2017.08.046>
- Mendoza, K. K. R., Zúñiga, L. H. P., & García, A. Y. L. (2024). Validity in evaluation: where is the argument-based approach heading? *Sapienza: International Journal of Interdisciplinary Studies*, 5(3), e24048-e24048. <https://doi.org/10.51798/sijis.v5i3.792>
- Messick, S. J. (2013). Alternative modes of assessment, uniform standards of validity. In *Beyond multiple choice* (pp. 59-74). Psychology Press.
- Metibemu, M. A., Orowole, K., & Adewale, J. G. (2024). Assessing the Functionality of National Examinations Council Physics Test Items Among Northern and Southern Candidates in Nigeria. *Journal of Educational Research and Development*, 11, 61-70.
- Mohammed, U., Ojo, E. O., & Mustapha, Y. A. (2024). Quality Educational Testing/Assessment: A Vital Consideration in Enhancing Sustainable Development in Nigeria. *Journal of Education*, 9(1), 127-133.
- Mohan, R. (2023). *Measurement, evaluation and assessment in education*. PHI Learning Pvt. Ltd.
- Moore, J. D., & Paris, C. L. (2019, October). Constructing coherent text using rhetorical relations. In *10th Annual Conference Cognitive Science Society* (pp. 637-643). Psychology Press.
- Obiakor, T. E. (2024). Language of instruction policy in Nigeria: Assessing implementation and literacy achievement in a multilingual environment. *International Journal of Educational Development*, 109, 103108. <https://doi.org/10.1016/j.ijedudev.2024.103108>
- Obidiegwu, U. J., & Chieke, J. C. (2021). Determination of The Adequacy of Programme Objectives in The NCECE 2012 Minimum Standards For Adult and Non-Formal Education In Nigeria. *African Journal of Educational Management, Teaching and Entrepreneurship Studies*, 3(1), 27-35.
- Obilor, E. I., & Akwawo, U. K. (2024). Influence of Frequent Testing on the Academic Performance of Students in Public Senior Secondary Schools in Akwa-Ibom State. *International Journal of Innovative Social & Science Education Research* 12(2), 91-107
- Obilor, E. I., & Omeke, K. S. (2024). Use of Test Blueprint in Improving Teachers' Test Construction Skills for Quality Assessment. *Journal of Educational Research and Development*, 11, 127-139.
- Ogunyinka, E. K., Okeke, T. I., & Adedoyin, R. C. (2015). Teacher education and development in Nigeria: An analysis of reforms, challenges and prospects. *Education Journal*, 4(3), 111-122. <https://doi.org/10.11648/j.edu.20150403.14>

- Okeke, R. C., & Chukwudebelu, C. B. (2024). Educational policy and national development in Nigeria. *International Journal of Education and Evaluation*, 10(4), 268-274.
- Oladele, B. K. (2021). *Comparison of Secondary School Students' Ability in Mathematics Constructed-Response Items Under Classical Test And Item Response Measurement Theories In The Ibadan Metropolis, Nigeria* (Doctoral Dissertation, University of Ibadan, Nigeria).
- Olaniyan, F. A. (2021). *School Management Effectiveness and Academic Performance of Selected Private Second Schools in Ilorin Metropolis* (Master's thesis, Kwara State University (Nigeria)).
- Oloba, P. B. (2023). *Stakeholders' Experiences and Perceptions of System-Wide Change in the Nigerian Basic Education School System*. University of Johannesburg (South Africa).
- Omachi, D. (2019). *Fundamentals of Educational Measurement and Evaluation*. Angusco Nigeria Enterprise.
- Onyeabo, E. A. (2021). *The Curse of Regulatory Transposition in Petroleum Governance and Energy Transition: A Nigerian Case Study* (Doctoral dissertation, The Chinese University of Hong Kong (Hong Kong)).
- Opposs, D., Baird, J. A., Chankseliani, M., Stobart, G., Kaushik, A., McManus, H., & Johnson, D. (2020). Governance structure and standard setting in educational assessment. *Assessment in Education: Principles, Policy & Practice*, 27(2), 192-214. <https://doi.org/10.1080/0969594X.2020.1730766>
- Phakiti, A., & Isaacs, T. (2021). Classroom assessment and validity: Psychometric and edumetric approaches. *European Journal of Applied Linguistics and TEFL*, 10(1), 3-24.
- Raji, M. O. (2023). *Conceptions and Practices of Assessment: Case Studies of Nigerian Teacher Educators* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- Raji, M. O., & Baidoo-Anu, D. (2025). Socioculturally Responsive Post-secondary Entrance Examination: Implications for Equitable Assessment Design in Sub-Saharan Africa. In *Socioculturally Responsive Assessment* (pp. 399-414). Routledge. <https://doi.org/10.4324/9781003435105-22>
- Singh, P., Thomas, S., & Harris, J. (2013). Recontextualising policy discourses: A Bernsteinian perspective on policy interpretation, translation, enactment. *Journal of Education Policy*, 28(4), 465-480. <https://doi.org/10.1080/02680939.2013.770554>
- Stone, J. E. (2021). *Self-determination theory and MOOC enrollment motivation: validation of the online learning enrollment intentions scale* (Doctoral dissertation, Oklahoma State University).
- Tobih, D. O., Ayanwale, M. A., Ajayi, O. A., & Bolaji, M. V. (2023). The use of measurement frameworks to explore the qualities of test items. *International journal of evaluation and research in education*, 12(2), 914-923. <https://doi.org/10.11591/ijere.v12i2.23747>
- Tobin, M., Nugroho, D., & Lietz, P. (2016). Large-scale assessments of students' learning and education policy: Synthesising evidence across world regions. *Research Papers in Education*, 31(5), 578-594. <https://doi.org/10.1080/02671522.2016.1225353>

- Ufondu, C. C., Alonta, G. C., & Amobi, S. C. (2023). Determining the Validity and Reliability of Business Studies Achievement Test for Senior Secondary Schools in Nigeria. *Futurity Education*, 3(3), 132-154.
- Yashim, A. U., & Jibrin, S. K. (2020). Continuous Assessment in Education: Problems and Prospects. *Journal of Educational Assessment and Pedagogical Process (JEAPP)* 1(1).
- Yuldashevna, E. F. (2025). Assessments For Students: Types, Purposes and Best Practices. *Educator Insights: Journal of Teaching Theory and Practice*, 1(4), 304-310.