DETERMINING THE CONSTRUCTS MEASURED BY THE WEST AFRICAN SENIOR SCHOOL CERTIFICATE EXAMINATION (WASSCE): A STUDY OF CONSTRUCT VALIDITY AND RELIABILITY

[*1]Innocent Ewaoche OCHEOLA, [2]Chinyere Jennifer UCHE-UDAH, [1]Dolly Ogo OKPARA, [3]Blessing Ene EJEGWA and [4]Chukwukaelo Chigbo EZIAMAKA

[1]Department of Educational Foundation, Federal College of Education Technical ISU, Ebonyi State.
[2]Department of Business Education, Federal College of Education Technical ISU, Ebonyi State.
[3]Department of Integrated Science, Federal College of Education Technical ISU, Ebonyi State.
[4]Department of Physics Education, Federal College of Education Technical ISU, Ebonyi State
[*]Corresponding Author: innocentewaoche@gmail.com

Abstract

This study examined the construct validity and reliability of the West African Senior School Certificate Examination (WASSCE) administered by the West African Examinations Council. Specifically, it investigated the extent to which WASSCE measures intended cognitive constructs through content representation and internal structure evidence. The study adopted an instrumentation and validation research design within a mixed-method framework. A total of 1,125 examination items from Mathematics, English Language, and Biology (2019–2023) were analyzed using Bloom's revised taxonomy, while student response data from 1,240 SS3 students were subjected to Exploratory Factor Analysis (EFA) and reliability analysis. Findings revealed that 85% of items assessed lower-order cognitive skills, while higher-order thinking was minimally represented. EFA results showed a dominant general factor accounting for 48% of variance, indicating weak multidimensional structure. Reliability coefficients ranged from .79 to .84, indicating acceptable internal consistency. The study concludes that while WASSCE is reliable, its construct validity is limited due to construct underrepresentation and weak cognitive differentiation. It is recommended that WAEC redesign test blueprints to strengthen higher-order cognitive assessment and adopt construct-based validation frameworks.

Keywords: Construct validity, reliability, WASSCE, cognitive levels, factor analysis, educational measurement

Introduction

Large-scale examinations are critical tools for certification and educational progression. In West Africa, the WASSCE, conducted by the West African Examinations Council, serves as the primary gateway to tertiary education. The validity of interpretations derived from such examinations depends on the extent to which they measure intended constructs such as knowledge, comprehension, and higher-order thinking skills.

Validity refers to the degree to which evidence supports the interpretation of test scores, while reliability concerns score consistency (AERA, APA, & NCME, 2014). Construct validity is central in educational measurement because it determines whether a test truly measures intended psychological or cognitive attributes (Messick, 1989). In achievement tests, constructs are represented through cognitive domains such as remembering, understanding, application, analysis, evaluation, and creation (Anderson & Krathwohl, 2001).

Despite the importance of WASSCE, concerns persist regarding its alignment with 21st-century competencies. Previous studies suggest an overemphasis on lower-order cognitive skills and limited structural validation of examination scores (Oladipo & Yusuf, 2019; Akintunde & Adebiyi, 2022). However, few studies integrate both content and structural evidence of validity within a unified framework.

This study therefore examines the constructs measured by WASSCE by analyzing cognitive representation, internal structure, and reliability of examination scores.

Literature Review

Construct Validity in Educational Assessment

Construct validity refers to the degree to which a test measures theoretical constructs it intends to measure (Kane, 2013). It is established through multiple sources of evidence including content, internal structure, and relationships with other variables (AERA et al., 2014).

Cognitive Structure of Achievement Tests

Bloom's revised taxonomy classifies cognitive processes into six levels, ranging from remembering to creating (Anderson & Krathwohl, 2001). Studies have shown that many standardized examinations overemphasize lower-order cognitive skills (Stobart, 2018).

Internal Structure of Large-Scale Examinations

Factor analytic studies often reveal that achievement tests exhibit a dominant general factor, suggesting unidimensionality (Embretson & Reise, 2000). In African contexts, similar findings have been reported in national examinations (Njoroge & Kagema, 2019).

Synthesis and Research Gap

Existing studies (Mensah & Boateng, 2021; Banjo & Smith, 2023) show that WASSCE tends to emphasize lower cognitive levels and may exhibit weak multidimensionality. However, most studies focus on either content or factor structure independently. Few integrate both evidence types to provide a comprehensive construct validity argument. This study fills this gap by combining cognitive analysis and factor analysis within a unified validation framework.

Methodology

The study adopted an instrumentation and validation research design within a mixed-method approach (Creswell & Creswell, 2018). This design is appropriate because validation requires multiple sources of evidence, including content and structural analysis (Messick, 1989).

The population consisted of all WASSCE examination items and candidates from 2019 to 2023 across WAEC member countries. However, the study focused on Nigeria due to its large candidate population. A census of 1,125 examination items in Mathematics, English Language, and Biology was used.

Additionally, 1,240 SS3 students were selected using multistage sampling for response analysis. Instruments included WASSCE papers, student response data, and a cognitive classification matrix based on Bloom's taxonomy. Validity of instruments was ensured through expert review, while reliability was assessed using Cohen's Kappa and Cronbach's alpha.

Data were analyzed using content analysis, Exploratory Factor Analysis (Principal Axis Factoring with Varimax rotation), and reliability analysis. EFA was used to determine whether WASSCE is unidimensional or multidimensional in structure (Kline, 2016). Findings from content and structural analyses were integrated to evaluate construct validity.

Results and Discussion

Cognitive Construct Representation

The distribution of WASSCE items across cognitive levels is presented in Table 1 below.

Table 1: Distribution of WASSCE Items by Cognitive Level (2019–2023)

Cognitive Level	Frequency	Percentage (%)
Remember	410	36.4
Understand	305	27.1
Apply	240	21.3
Analyze	95	8.4
Evaluate	45	4.0
Create	30	2.7
Total	1,125	100

Interpretation: Table 1 shows that 84.8% of items fall within lower and mid level cognitive categories (remember, understand,apply), while only 15.1% represent higher order thinking skills.

Table 2: Eigenvalues and Variance Explained by Extracted Factors

Factor	Eigenvalue	%Variance	cumulative%
Factor 1	4.80	48.0	48.0
Factor 2	1.30	13.0	61.0
Factor 3	0.90	9.0	70.0

Interpretation: Table 2 indicates that Factor 1 explains the largest proportion of variance (48%), suggesting a dominant general factor. Subsequent factors contribute minimally, indicating weak multidimentionality.

Rotated Factor Loadings

Table 3: Sample Rotated Factor Loadings of Selected Items

Item	Factor 1	Factor 2	Factor 3
Item 1(math)	0. 72	0.21	0.10
Item 2(Eng)	0.68	0.18	0.14
Item 3(Bio)	0.75	0.20	0.09
Item 4	0.70	0.25	0.11
Item 5	0.73	0.19	0.12

Interpretation: Items load strongly on Factor 1, confirming the dominance of a general achievement construct.

Reliability Analysis

Table 4: Cronbach's Alpha Reliability Coefficients

Subject	Cronbach's Alpha(α)
Mathematics	0.82
English	0.79
Biology	0.84

Interpretation: All subjects demonstrate acceptable reliability ($\alpha \geq 0.70$), indicating consistent measurement across items

Integration of Content and Structural Evidence

Table 5: Comparison of intended VS Observed Construct Structure

Evidence Type	Expected (Content)	Observed (EFA)	Interpretation
Cognitive Levels	Multiple levels	Weak seperation	Partial mismatch
Higher-order skills	Present	Not distinct	Underrepresented
Construct structure	Multidimensional	largely unidimensional	General factor dominance

Interpretation: Table 5 shows a mismatch between intended cognitive constructs and observed factor structure, indicating limitations in construct validity.

Summary of Findings

Based on the analysis of examination items and student response data for the West African Senior School Certificate Examination (WASSCE) conducted by the West African Examinations Council, the following detailed findings emerged:

First, the analysis of cognitive construct representation revealed a significant imbalance in the distribution of examination items across cognitive levels. Specifically, 84.8% of the total items analyzed were concentrated within the lower and mid-level cognitive domains of remembering (36.4%), understanding (27.1%), and application (21.3%). In contrast, only 15.1% of the items assessed higher-order cognitive processes, including analysis (8.4%), evaluation (4.0%), and creation (2.7%). This finding indicates a strong dominance of lower-order cognitive demand in WASSCE, suggesting that the examination places greater emphasis on recall and basic comprehension rather than critical thinking and problem-solving skills.

Second, the internal structure analysis using Exploratory Factor Analysis (EFA) revealed that WASSCE scores are largely explained by a single dominant factor. The first factor accounted for approximately 48% of the total variance, while subsequent factors contributed substantially less and did not exhibit clear interpretability in terms of distinct cognitive constructs. This pattern indicates that the examination operates predominantly as a

unidimensional measure of general academic achievement, rather than a multidimensional assessment capturing separate cognitive domains such as reasoning, analysis, or evaluation.

Third, the examination demonstrated acceptable levels of internal consistency reliability across all subjects analyzed. Cronbach's alpha coefficients ranged from 0.79 to 0.84 for Mathematics, English Language, and Biology, indicating that the items consistently measure a common underlying construct. However, while these reliability indices confirm measurement consistency, they do not provide evidence of adequate construct representation or differentiation.

Fourth, the integration of content-based and structural evidence revealed a notable misalignment between the intended and observed constructs measured by WASSCE. Although the content analysis suggests that higher-order cognitive skills are included in the examination framework, the factor analytic results indicate that these skills are not distinctly reflected in the structure of student performance data. This discrepancy points to construct underrepresentation, where certain intended cognitive domains are insufficiently captured in actual measurement.

Fifth, the combined evidence from content distribution and internal structure suggests that WASSCE functions more as a general achievement test than a construct-differentiated assessment instrument. The dominance of lower-order cognitive items, coupled with a strong general factor, implies that the examination does not effectively discriminate between different levels of cognitive ability.

Overall, the findings provide converging evidence that, while WASSCE is reliable, it exhibits limitations in construct validity, particularly in terms of higher-order cognitive representation and multidimensional measurement structure

Conclusion

This study set out to determine the extent to which the West African Senior School Certificate Examination (WASSCE), administered by the West African Examinations Council, measures intended cognitive constructs through an examination of content representation, internal structure, and reliability of test scores.

Based on the empirical evidence generated, it is concluded that WASSCE demonstrates acceptable reliability but limited construct validity. The analysis of cognitive construct representation revealed a substantial imbalance in item distribution, with the overwhelming majority of items concentrated at lower cognitive levels remembering, understanding, and application while higher-order cognitive processes such as analysis, evaluation, and creation were minimally represented. This indicates that the examination is predominantly designed to assess basic knowledge acquisition and routine problem-solving rather than advanced cognitive skills.

Furthermore, the internal structure analysis using Exploratory Factor Analysis showed that student performance is largely explained by a single dominant general factor, with weak and poorly defined secondary dimensions. This suggests that WASSCE operates primarily as a unidimensional measure of general academic achievement, rather than a multidimensional assessment capable of distinguishing between different cognitive constructs. The implication is that even when higher-order items are included, they do not function distinctly enough to form separate measurable constructs in the performance data.

Although the reliability coefficients obtained for the examination were within acceptable limits, indicating internal consistency of measurement, this alone does not guarantee validity. Reliability without adequate construct representation implies that the examination consistently measures a construct that may not fully align with its intended cognitive objectives.

The integration of content and structural evidence further revealed a misalignment between intended and observed constructs. While the examination framework suggests coverage of multiple cognitive domains, the empirical structure derived from student responses does not reflect this multidimensionality. This points to construct underrepresentation, where important aspects of cognitive functioning particularly higher-order thinking are insufficiently captured.

In summary, the study concludes that WASSCE functions more effectively as a general achievement assessment rather than a construct-differentiated instrument. While it provides consistent measurement outcomes, its ability

to validly represent and distinguish among diverse cognitive constructs is limited. This has important implications for the interpretation of examination results and their use in high-stakes educational decisions.

Recommendations

In light of the findings and conclusions of this study, the following recommendations are made:

1. The West African Examinations Council should undertake a comprehensive review of the WASSCE test blueprint to ensure a more balanced distribution of cognitive levels, with deliberate inclusion of higher-order thinking skills such as analysis, evaluation, and creation.
2. Test developers should adopt construct-centered item development approaches, ensuring that each item is explicitly designed to measure clearly defined cognitive constructs, thereby improving the alignment between intended and observed measurement structures.
3. Regular psychometric validation studies should be institutionalized, including the use of advanced techniques such as Confirmatory Factor Analysis (CFA) and Item Response Theory (IRT), to continuously evaluate and refine the construct validity of WASSCE.
4. Capacity-building programmes should be organized for item writers and examiners to strengthen their competence in developing higher-order cognitive assessment items, particularly those that assess critical thinking, reasoning, and problem-solving skills.
5. Curriculum planners and teachers should align classroom instruction with higher-order cognitive objectives to ensure that students are adequately prepared for more cognitively demanding assessments.
6. Future research should expand the scope of validation studies to include additional subjects and cross-country samples within WAEC member states, in order to enhance the generalizability of findings and provide a more comprehensive understanding of the examination's measurement properties.

References

AERA, APA, & NCME. (2014). Standards for educational and psychological testing. *American Educational Research Association.*

Adeoye, A. Y., & Adebayo, A. D. (2021). *Dimensionality analysis of national examination scores: Evidence from Nigeria. Journal of Educational Measurement and Evaluation, 9(2), 45–63.*

Akintunde, O. O., & Adebiyi, S. O. (2022). *Cognitive demand analysis of WASSCE mathematics items. African Journal of Mathematics Education, 15(1), 78–91.*

Alabi, O. O., & Adeyemi, A. A. (2022). From curriculum intent to classroom practice: *The role of examination demands in shaping instruction in Nigeria. Journal of Curriculum Studies, 14(4), 211–230.*

Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: *A revision of Bloom's taxonomy of educational objectives. Longman.*

Banjo, T. L., & Smith, J. E. (2023). Factor structure of WASSCE science scores: *An exploratory assessment. International Journal of Science Education and Measurement, 10(3), 33–50*.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: *Principles, Policy & Practice, 5(1), 7–74.*

Creswell, J. W., & Creswell, J. D. (2018). Research design. Sage.

ECOWAS Commission. (2019). Regional education policy framework. ECOWAS Publications.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists. Lawrence Erlbaum Associates.*

Federal Ministry of Education (Nigeria). (2013). National policy on education (6th ed.). *NERDC Press.*

Ghazali, N., & Cherif, M. (2020). Construct representation in large-scale assessments: *Challenges and opportunities. Assessment Today, 4(1), 55–74.*

Hambleton, R. K., & Pitoniak, M. J. (2006). Large-scale assessment in education. *Wiley.*

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50(1), 1–73.*

Kline, R. (2016). Principles and practice of structural equation modeling. Guilford.

Mensah, P., & Boateng, O. (2021). Cognitive demand and item quality in WASSCE English papers. *Journal of Language Assessment and Education, 7(2), 117–134.*

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., *pp. 13–103). Macmillan.*

Millsap, R. E. (2011). Statistical approaches to measurement invariance. *Routledge.*

Mushi, S., & Kijakazi, J. (2022). Cognitive demand analysis of Tanzanian secondary school certificates. *African Educational Assessment Review, 8(1), 61–78.*

Njoroge, R. W., & Kagema, J. M. (2019). Factor structure of Kenya Certificate of Secondary Education: A psychometric evaluation. *East African Journal of Education Research, 5(3), 24–40.*

Okoro, C. C. (2021). Curriculum implementation challenges in Nigerian secondary schools: Teachers' perspectives. *Nigerian Journal of Educational Practice, 12(1), 9–25.*

Okonkwo, I. A., & Eze, M. (2021). Assessment practices in Nigerian secondary schools: A critique of national examinations. *Journal of Educational Development, 3(4), 12–29.*

Oladipo, F. A., & Yusuf, M. O. (2019). Cognitive levels of examination items in Nigerian secondary schools. *International Journal of Educational Research, 10(2), 86–104*.

Popham, W. J. (2008). Classroom assessment: What teachers need to know (5th ed.). *Pearson.*

Popham, W. J. (2017). Test better, teach better: *The instructional role of assessment. ASCD*.

Pretorius, E., & Ford, T. (2020). Content demands of national exams in South Africa: *A cognitive analysis. South African Journal of Educational Assessment, 6(1), 72–88.*

Stobart, G. (2018). Assessment: A critical perspective. *Routledge.*

Webb, N. L. (2019). *Alignment and cognitive complexity in assessment. Springer.*

Webb, N. L., & Jones, J. F. (2018). Assessing higher-order thinking in standardized *testing regimes. Educational Measurement Quarterly, 15(2), 145–168.*