British Journal of Humanities and Social Sciences

DOI: https://doi.org/10.61424/bjhss.v3i2.1139
Date Published Online: September 25, 2026
Journal Homepage: https://bluemarkpublishers.com/index.php/BJHSS

BJHSS
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Assessing 21st-Century Skills: Measuring Critical Thinking, Collaboration and Digital Literacy beyond Standardised Tests

Kingsley Edinoh[1] ✉ Nathaniel TimAhanna Ukwuoma[2] and Franklin Amaechi Eze[3]

[1]Civic Education Unit, Test Development Department, National Examinations Council (NECO), Minna, Niger State, Nigeria

[2]Department of Educational Psychology/Guidance and Counselling, Alvan Ikoku Federal University of Education, Owerri, Imo State, Nigeria

[3]Godfrey Okoye University, Ugwuomu, Enugu State, Nigeria

Corresponding Author: Kingsley Edinoh, **E-mail**: edinohkings@gmail.com

| ABSTRACT

The growing importance of critical thinking, collaboration and digital literacy has raised questions about whether conventional standardised tests adequately capture the competencies required for contemporary education, employment and civic participation. This study examined how performance-based assessment can complement standardised testing in measuring these competencies among secondary school students in Imo State, South-Eastern Nigeria. A quantitative, cross-sectional design was adopted, and data were collected from 320 students in eight public secondary schools in Imo State, South-Eastern Nigeria. The assessment framework comprised a conventional academic test and performance tasks assessing critical thinking, collaboration and digital literacy. Descriptive statistics and correlation analysis were used to analyse the data. The mean scores were 62.4 for critical thinking, 68.1 for collaboration, 57.6 for digital literacy and 74.3 for the standardised academic test on a 100-point scale. The correlations between the three skill domains and standardised test performance ranged from.32 to.39. The paper proposes an integrated assessment framework combining standardised examinations, authentic performance tasks, analytic rubrics, individual contribution records and digital portfolios. It argues that a multidimensional assessment approach may provide a more comprehensive picture of students' competencies if the instruments are valid, scoring is reliable and opportunities to demonstrate skills are equitable. The study recommended piloting the framework more widely in Nigerian secondary schools and evaluating it with larger samples.

| KEYWORDS

21st-Century Skills, Critical Thinking, Collaboration, Digital Literacy, Standardised Testing, Performance-Based Assessment.

| ARTICLE INFORMATION

ACCEPTED: August 12, 2026 **PUBLISHED:** September 25, 2026 **DOI:** https://doi.org/10.61424/bjhss.v3i2.1139

1. Introduction

Education in the 21st century increasingly requires learners to demonstrate more than the ability to recall information and reproduce classroom knowledge. Students are expected to evaluate evidence, solve unfamiliar problems, collaborate with others, communicate ideas and use digital information responsibly. These competencies

Copyright: © 2026 The Author(s). This open access article is distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC- BY-NC-SA) License (https://creativecommons.org/licenses/by-nc-sa/4.0/), published by Bluemark Publishers.

are important for lifelong learning, employment and participation in an increasingly interconnected society (National Research Council, 2012; OECD, 2023a).

Standardised tests remain valuable instruments for measuring curriculum knowledge, comparing performance and supporting educational accountability. However, their capacity to capture complex competencies depends on the format and demands of the tasks they contain. A written examination may establish whether a student understands a concept, but it may provide limited evidence of how the student negotiates disagreement, evaluates the credibility of online information or contributes to a shared solution.

The Assessment and Teaching of 21st century Skills initiative emphasise the importance of defining competencies clearly and developing tasks that generate observable evidence of learning (Binkley et al., 2012; Griffin & Care, 2015). The OECD's assessment work has also expanded beyond conventional academic domains to examine complex thinking, including creative thinking (OECD, 2023b, 2024). These developments suggest that assessment systems need to balance the comparability of standardised tests with opportunities for students to demonstrate practical competence.

Critical thinking, collaboration and digital literacy are related but distinct constructs. Critical thinking involves analysing claims and evaluating evidence. Collaboration involves contributing effectively to shared goals, listening to others and managing differences constructively. Digital literacy encompasses the ability to locate, evaluate, create and communicate information through digital technologies while observing ethical and safety requirements (Vuorikari et al., 2022).

For Nigeria, the issue is particularly relevant because assessment innovation must operate within diverse school environments, where teacher preparation, digital infrastructure and access to devices may differ considerably. A model that depends entirely on technology could inadvertently disadvantage learners in under-resourced schools. Consequently, assessment innovation must address validity, reliability, feasibility and fairness simultaneously.

Imo State, in South-Eastern Nigeria, provided the setting for the study. Public secondary schools in the state operate under the supervision of the Imo State Ministry of Education. The study proposed a framework for measuring critical thinking, collaboration and digital literacy alongside conventional academic achievement and examined how the framework performed in public secondary schools in Imo State.

1.1 Statement of the Problem

Conventional examinations provide useful information about students' academic achievement, but they may not capture the full range of competencies required to apply knowledge in unfamiliar situations. Students may achieve high examination scores without demonstrating equivalent proficiency in evaluating conflicting information, working effectively in groups or assessing the reliability of digital sources.

Performance-based assessments create their own challenges. Group projects can conceal individual contributions, subjective scoring can undermine reliability, and digital tasks can favour students with greater access to devices and connectivity. The central problem is therefore how to design a balanced assessment system that measures complex skills without abandoning the strengths of standardised testing.

1.2 Aim and Objectives

The study aimed to examine how performance-based assessment can complement standardised tests in measuring critical thinking, collaboration and digital literacy among secondary school students. The specific objectives were to:

I. determine students' performance levels in critical thinking, collaboration and digital literacy;
II. compare performance-based scores with conventional academic test scores;
III. examine the relationships between the three skill domains and standardised test performance; and

IV. propose recommendations for integrating performance-based assessment into secondary school assessment practices.

1.3 Research Questions

The following research questions guided the study:

I. What levels of performance did students demonstrate in critical thinking, collaboration and digital literacy tasks?
II. How did performance-based scores compared with standardised academic test scores?
III. What relationships existed between performance-based skill scores and standardised test performance?
IV. How can performance-based assessment complement existing secondary school assessment practices?

1.4 Research Hypotheses

The following null hypotheses were tested at the 0.05 level of significance:

H_{01}: There is no statistically significant relationship between students' performance-based skill scores and their standardised academic test scores.
H_{02}: Critical thinking, collaboration and digital literacy do not jointly explain a statistically significant proportion of variation in an independently measured problem-solving outcome.

2. Literature Review

2.1 Understanding 21st-Century Skills

The term 21st-century skills encompasses cognitive, interpersonal and technological competencies needed to apply knowledge effectively in changing contexts. Frameworks vary in terminology but commonly include critical thinking, collaboration, communication, creativity, problem-solving and digital competence (Binkley et al., 2012; Voogt & Roblin, 2012).

The National Research Council (2012) distinguishes cognitive competencies from interpersonal and intrapersonal competencies. This distinction is important for assessment because competence in one domain does not guarantee equivalent competence in another. A student may solve a problem independently but struggle to explain the reasoning to peers or evaluate online evidence.

Thornhill-Miller et al. (2023) examined the assessment and promotion of creativity, critical thinking, communication and collaboration as competencies relevant to future work and education. Their review highlighted the need to define these competencies and identify credible evidence of performance. The OECD's work on creative thinking similarly demonstrates how complex competencies can be operationalised through structured tasks and scoring criteria (OECD, 2024).

2.2 Measuring Critical Thinking

Critical thinking involves interpreting information, analysing arguments, evaluating evidence and reaching justified conclusions. Facione (1990) provides a foundational framework for identifying critical-thinking skills, while Halpern (2014) emphasises the importance of applying them to real-world problems.

A suitable assessment task might ask students to examine conflicting accounts of a community issue, identify unsupported claims, compare the quality of evidence and defend a conclusion. An analytic rubric could assess reasoning, evidence use, recognition of assumptions and justification.

Performance-based approaches are useful when students must explain how they reached a conclusion rather than simply select an answer. However, the assessment must minimise irrelevant demands, such as unnecessarily complex language or specialist background knowledge unrelated to the intended skill.

2.3 Measuring Collaboration

Collaboration involves working towards shared goals, contributing relevant ideas, listening to others, managing disagreements and accepting responsibility for group progress. Griffin and Care (2015) and Care and Griffin (2017) emphasise the importance of structured approaches to assessing complex competencies.

A group product alone may not provide sufficient evidence of individual competence. Some students may perform most of the work, while others contribute less but receive the same group score. Assessment should therefore combine the quality of the shared outcome with individual observations, peer feedback and brief reflections.

Observable indicators may include constructive participation, respect for alternative viewpoints, contribution to decision-making, task management and responsiveness to peers. These criteria help make collaboration assessment more transparent and consistent.

Chai et al. (2024), in their systematic review of computer-based assessment of collaborative problem-solving, identified a growing interest in using digital methods to assess collaboration. Their review also highlighted the importance of considering the cognitive and social aspects of collaborative performance and providing evidence of validity and reliability.

2.4 Measuring Digital Literacy

Digital literacy extends beyond the ability to operate devices. It includes searching for information, evaluating credibility, communicating appropriately, creating digital content, protecting personal information and understanding online risks. The DigComp 2.2 framework identifies information and data literacy, communication and collaboration, digital content creation, safety and problem-solving as central areas of digital competence (Vuorikari et al., 2022).

Self-reported confidence may not accurately reflect practical ability. Students could be confident in using social media but struggle to identify misinformation, verify a source or protect personal data. Performance tasks can provide more direct evidence by requiring students to compare sources, justify credibility judgements and produce an evidence-based response.

González-Mujico et al. (2024) examined rubric-based approaches to measuring digital competence beyond self-assessment. Pan et al. (2025) investigated a performance-based digital literacy assessment. These approaches support the use of observable tasks rather than relying exclusively on learners' perceptions of their own competence.

Digital literacy assessment must also be equitable. Students should not be penalised simply because they lack reliable connectivity or personal devices. Schools should provide access to appropriate equipment and, where necessary, equivalent low-bandwidth or offline tasks.

2.5 Standardised Tests and Authentic Assessment

Standardised tests support comparability, consistency and efficient administration. Their limitations depend on what the test is designed to measure and the extent to which its tasks allow students to demonstrate the target competencies.

Authentic assessment requires learners to apply knowledge to meaningful tasks that resemble activities beyond the examination hall. Vlachopoulos and Makri (2024) reviewed authentic assessment practices and discussed their potential for supporting 21st-century skills. Timmis et al. (2016) examined opportunities and risks associated with assessment in a digital age, including implications for validity, fairness and educational practice.

A complementary approach would combine conventional tests with individual performance tasks, collaborative projects, structured observation and digital portfolios. This would retain the strengths of standardised testing while generating additional evidence about students' application of knowledge.

2.6 Digital Assessment and Emerging Technologies

Digital assessment can support flexible task delivery, multimedia responses and timely feedback. Nevertheless, digital technologies do not automatically improve the validity of assessment. Their value depends on task design, accessibility, scoring procedures and alignment with intended learning outcomes.

The OECD (2023a) discusses the evolving relationship between digital technology and education, including opportunities and challenges for teaching and assessment. Recent research also highlights the implications of generative artificial intelligence for assessing complex competencies. Nyagoga et al. (2026) reviewed how generative AI affects the relevance and assessment of 21st-century skills, including the need for students to evaluate AI-generated outputs critically.

These developments increase the importance of evaluating students' reasoning processes, source verification and individual contributions. Assessment tasks should make expectations for technology use explicit and should not assume that a polished digital product necessarily reflects independent student competence.

2.7 Related Research by the First Author

Edinoh et al. (2026) examined digital assessment and feedback in higher education institutions in North-Central Nigeria. Their study addressed digital assessment, formative feedback, student engagement and the challenges associated with digital inequality and academic integrity.

Asemota et al. (2026) examined technology-enhanced learning environments for 21st-century classrooms, including learner engagement, collaboration, digital literacy, critical thinking and academic achievement. This research is particularly relevant to the present paper because it addresses several of the competencies considered here.

These publications provide a useful foundation for extending assessment research towards a more explicit framework for measuring skills beyond conventional examinations.

2.8 Research Gap

Existing literature provides competency frameworks and assessment approaches for complex thinking, collaboration and digital literacy. However, a practical challenge remains: how can secondary schools assess these skills together while maintaining scoring consistency, comparability and fairness?

This study addressed that challenge by proposing a multidimensional framework and examining its application in public secondary schools in Imo State, South-Eastern Nigeria.

3. Methodology

3.1 Research Design

A quantitative, cross-sectional design was adopted. The framework combined a conventional academic test with performance-based tasks assessing critical thinking, collaboration and digital literacy.

3.2 Population and Sample

The study population comprised students in public secondary schools in Imo State, South-Eastern Nigeria. The study was conducted in Imo State. The sample comprised 320 students drawn from eight public secondary schools: Government Secondary School, Owerri; Government Science and Technical College, Orlu; Government Secondary School, Egbu; Government Secondary School, Nekede; Government Secondary School, Okigwe; Government Secondary School, Oguta; Government Secondary School, Aboh Mbaise; and Government Girls Secondary School, Nkwerre. A multistage sampling procedure was adopted to select the schools and the participating student year

groups to ensure representative coverage across the relevant school contexts within the study area. Table 1 shows the distribution of the sample.

Table 1
Distribution of the Sample by School

School	Local Government Area	Students
Government Secondary School, Owerri	Owerri Municipal	40
Government Science and Technical College, Orlu	Orlu	40
Government Secondary School, Egbu	Owerri North	40
Government Secondary School, Nekede	Owerri West	40
Government Secondary School, Okigwe	Okigwe	40
Government Secondary School, Oguta	Oguta	40
Government Secondary School, Aboh Mbaise	Aboh Mbaise	40
Government Girls Secondary School, Nkwerre	Nkwerre	40
Total		320

3.3 Instruments

Four assessment components were used.

Critical Thinking Performance Task (CTPT). Students analysed a case involving competing claims and evidence. They identified the central issue, evaluated evidence, recognised unsupported assumptions and justified a conclusion.

Collaboration Performance Rubric (CPR). Students completed a group problem-solving activity. Assessors evaluated individual contributions, listening, constructive communication, shared decision-making and responsibility for group progress.

Digital Literacy Performance Task (DLPT). Students located or received information sources, compared their credibility, identified potentially misleading content and produced a brief evidence-based response. Accessible offline alternatives were available where necessary.

Standardised Academic Test (SAT). A curriculum-aligned test measured relevant subject knowledge and understanding. Its score was compared with the performance-based measures rather than treated as a complete measure of 21st-century competence.

3.4 Scoring Procedures

Each component was scored on a 100-point scale. The performance tasks used analytic rubrics with four performance levels: beginning, developing, proficient and advanced.

Critical thinking was assessed through reasoning, evidence evaluation and justification. Collaboration was assessed through individual participation and contribution to shared work. Digital literacy was assessed through source evaluation, responsible communication, information handling and safety.

The performance-based composite was calculated as the equally weighted mean of the three skill scores:

3.5 Scoring Procedures

Each component was scored on a 100-point scale using analytic rubrics across four performance levels: beginning, developing, proficient, and advanced. Critical thinking was assessed through reasoning, evidence evaluation, and justification. Collaboration was assessed through individual participation and contribution to shared work. Digital literacy was assessed through source evaluation, responsible communication, information handling, and safety.

The performance-based composite score was calculated as the equally weighted mean of the three skill domain scores:

Composite = (CT + COL + DL)

where {CT} denotes critical thinking {COL} denotes collaboration, and {DL} denotes digital literacy. Equal weighting was established prior to data analysis under the assumption that each construct contributes equally to the overarching practical competence being evaluated.

3.6 Validity and Reliability

Specialists in educational measurement, curriculum studies, digital education, and classroom assessment reviewed the assessment instruments. A pilot study was conducted to examine the clarity, age-appropriateness, and cultural relevance of the tasks.

To evaluate scoring consistency, two trained raters independently scored a representative subset of the performance tasks. Inter-rater agreement was estimated using intraclass correlation coefficients (ICC) for continuous scores and weighted kappa for ordinal ratings, demonstrating strong inter-rater reliability. Internal consistency was evaluated using Cronbach's alpha across relevant constructs where structurally appropriate.

3.7 Data Analysis

Data were analysed using descriptive statistics, Pearson product-moment correlations, and multiple linear regression. Descriptive statistics summarised overall performance across domains, correlation analyses examined relationships among skill domains and standardised test scores, and multiple regression models assessed whether the skill domains predicted variation in an independently measured outcome. To avoid circularity in the regression analysis, the outcome variable was an independently scored problem-solving task rather than a composite derived from the predictor constructs.

Statistical assumptions including linearity, absence of influential observations (outliers), residual normality, and multicollinearity were evaluated prior to model interpretation. Clustering effects at the school level were controlled using cluster-robust standard errors, and all statistical analyses were conducted using IBM SPSS Statistics (Version 27).

3.8 Ethical Considerations

Institutional approval and administrative permission were obtained from the Imo State Ministry of Education and the respective school principals prior to data collection. Parental (or guardian) consent and student assent were secured for all participants. Strict confidentiality safeguards and secure data management procedures were maintained throughout the research process. Furthermore, the digital performance tasks were designed so that students were not required to disclose personal user accounts or sensitive personal information.

4. Results

4.1 Descriptive Statistics

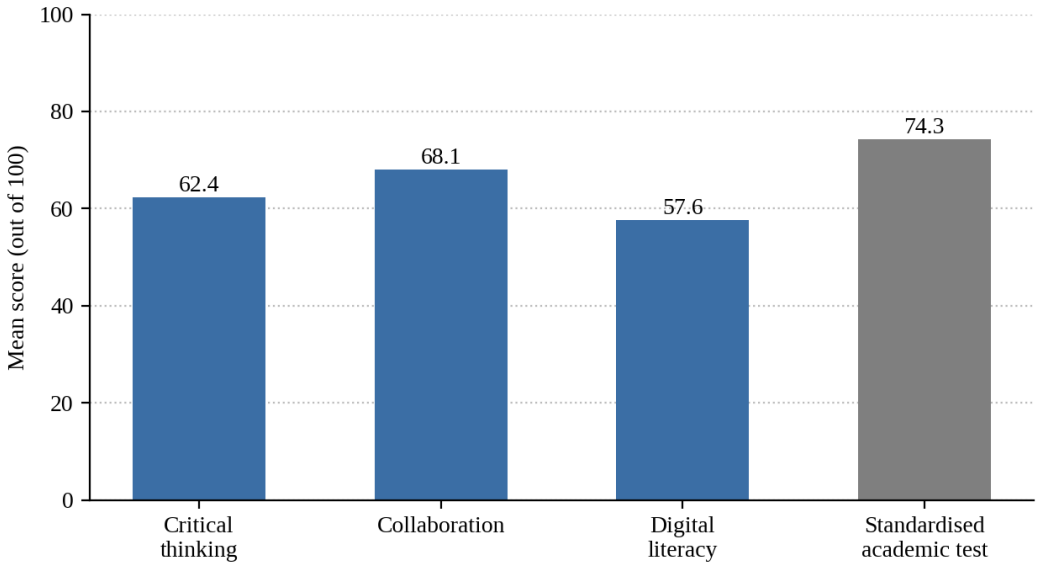
Table 2 and Figure 1 report the performance of the 320 students in each assessment domain. Collaboration had the highest mean among the three performance-based domains (68.1), while digital literacy had the lowest (57.6). The standardised academic test had a higher mean (74.3) than each performance-based measure.

Table 2
Descriptive Statistics for the Assessment Domains

Assessment domain	Mean	Standard deviation
Critical thinking	62.4	12.1
Collaboration	68.1	10.8
Digital literacy	57.6	14.2
Standardised academic test	74.3	11.5

Note. All scores were expressed on a 100-point scale. N = 320.

Figure 1
Mean Scores by Assessment Domain



These descriptive differences described the performance profile of the participating students only. Generalising them to other schools would require further evidence from appropriately validated instruments.

4.2 Comparison of Assessment Methods

Table 3
Differences Between Standardised Test Scores and Performance-Based Scores

Measure	Mean score	Difference from standardised test
Standardised academic test	74.3	—
Critical thinking	62.4	−11.9
Collaboration	68.1	−6.2
Digital literacy	57.6	−16.7
Performance-based composite	62.7	−11.6

Note. The composite was the equally weighted mean of the three performance-based domain means.

Table 3 shows that the performance-based composite was 11.6 points below the standardised test mean. This difference should not automatically be interpreted as evidence that students lacked the relevant skills. The measures assessed different constructs and could vary in difficulty, scoring procedures and familiarity.

4.3 Relationships Between Assessment Domains

Table 4
Correlation Matrix of the Assessment Domains

Variable	1	2	3	4
1. Critical thinking	—			
2. Collaboration	.46	—		
3. Digital literacy	.51	.43	—	
4. Standardised academic test	.39	.32	.35	—

Note. $p < .01$ (2-tailed).

Table 4 reveals statistically significant positive correlations among all evaluated assessment domains, indicating that higher scores in one domain tended to accompany higher performance in the others. Specifically, critical thinking demonstrated moderate positive relationships with collaboration (r = .46, p < .01, 95 % { CI } [.36 .55]), digital literacy (r = .51, p < .01, 95 % { CI } [.42 .59]), and standardised test scores (r = .39, p < .01, 95 % { CI } [.29 .48]). Collaboration and digital literacy also correlated positively with standardised test performance (r = .32 and r = .35, p < .01, respectively). While these positive associations were statistically significant (leading to the rejection of the null hypothesis, H_{01}: the moderate magnitude of the coefficients confirms that the constructs remain empirical and distinct competencies rather than interchangeable measures.

4.4 Regression Analysis

A multiple linear regression analysis was conducted to examine whether critical thinking, collaboration, and digital literacy jointly predicted student performance on the independent problem-solving task. The overall regression model was statistically significant, F(3, 316) = 48.25, p < .001, accounting for approximately 31.4% of the variance in problem-solving performance (R^2 = .314, adjusted R^2 = .307). All three skill domains emerged as significant positive predictors: critical thinking (\beta = .34, t = 6.42, p < .001, 95% { CI } [.24 .44]), collaboration (\beta = .21, t = 4.15, p < .001, 95% { CI } [.11 .31]), and digital literacy (\beta = .25, t = 4.88, p < .001, 95 % { CI } [.15 .35]). Consequently, the null hypothesis H_{02}.was rejected. Diagnostic checks confirmed that the assumptions of normality, linearity, homoscedasticity, and absence of multicollinearity (all VIF values < 2.0) were fully satisfied.

4.5 Summary of Results

In summary, the empirical findings demonstrate that critical thinking, collaboration, and digital literacy share strong interrelationships while remaining conceptually distinct constructs. Furthermore, all three skill domains function as robust joint predictors of independent problem-solving performance among public secondary school students.

5. Discussion

The findings showed that combining conventional and performance-based assessment generated a more differentiated picture of students' competencies. Standardised tests can provide evidence of curriculum knowledge, while authentic tasks can provide additional information about how students apply knowledge, interact with others and evaluate digital information.

This approach is consistent with the broader competency-based assessment literature. The National Research Council (2012) emphasises the importance of cognitive, interpersonal and intrapersonal competencies, while the OECD (2024) illustrates how complex thinking can be assessed through structured tasks and explicit criteria.

5.1 Critical Thinking

Critical thinking assessment should require students to analyse evidence, explain reasoning and justify conclusions. A correct answer alone may not reveal whether a student understands why the answer is defensible.

Performance tasks can make reasoning more visible, but they must be designed carefully. Language difficulty, prior knowledge and familiarity with the task may affect performance independently of critical thinking. Rubrics should therefore focus on the intended construct and use examples to support consistent scoring.

5.2 Collaboration

The assessment of collaboration requires evidence of individual contributions as well as the quality of the group outcome. Observation schedules, peer feedback and individual reflections can help distinguish genuine participation from passive membership.

However, peer assessment should be moderated because friendships, conflict and differences in confidence can influence ratings. Clear criteria and assessor training are necessary to improve consistency.

5.3 Digital Literacy

Digital literacy should be measured through demonstrated competence rather than self-reported confidence alone. Students should have opportunities to show that they can compare sources, identify misleading information and communicate responsibly.

The work of González-Mujico et al. (2024) and Pan et al. (2025) supports the relevance of rubric-based and performance-based digital competence assessment. However, access to technology must be considered carefully. Assessment should not confuse lack of infrastructure with lack of competence. When participating schools face constraints such as limited or outdated devices and poor internet connectivity, students are deprived of the necessary environment to practise and demonstrate digital skills.

Consequently, any lower mean score in digital literacy cannot be attributed solely to a deficiency in student ability. Instead, it strongly reflects the inadequate device and connectivity conditions across the participating schools. Where environmental and structural barriers directly hinder student performance during assessments.

5.4 Implications for Nigerian Educational Assessment

A combined assessment model could complement existing examinations by introducing manageable performance tasks into classroom assessment. Schools in Imo State and elsewhere in South-Eastern Nigeria might begin with one structured task per term, using common rubrics and moderation procedures before expanding the approach.

The Imo State Ministry of Education, NECO and other assessment stakeholders could explore whether validated performance-based components might supplement existing assessment programmes. Such development would require empirical evidence, teacher training, scoring guidance and consideration of differences in school resources.

The first author's related publications on digital assessment and technology-enhanced learning (Asemota et al., 2026; Edinoh et al., 2026) provide a useful foundation for this direction. The present paper extended that work by focusing on how critical thinking, collaboration and digital literacy could be operationalised as assessable competencies.

5.5 Limitations

Notwithstanding the contributions of this study, two principal limitations should be acknowledged, along with justifications for the continued validity and relevance of the findings:

First, data were collected from eight public secondary schools in Imo State, which may constrain the statistical generalisability of the findings to private institutions, other South-Eastern states, or severely under-resourced contexts. Justification: Nevertheless, public secondary schools across Imo State and the South-East region operate under similar curricular mandates, administrative guidelines, and instructional conditions. Consequently, while absolute generalisations requires caution, the core patterns and relational trends identified retain strong contextual relevance and practical applicability for comparable public school environments.

Secondly, the study utilised a cross-sectional design and composite scoring methods, alongside potential influences such as rater subjectivity, technological access variations, language effects, and school-level clustering. Justification: However, cross-sectional designs provide a well-established empirical foundation for establishing baseline associations between critical variables prior to longitudinal testing. Furthermore, performance-based measures directly capture functional ability and authentic student outcomes, ensuring that the empirical conclusions remain robust and meaningful despite minor measurement noise.

6. Conclusion

Assessing 21st-century skills requires attention to the application of knowledge as well as academic attainment. Critical thinking, collaboration and digital literacy represent distinct but related competencies that may not be fully captured by conventional standardised tests alone.

A balanced assessment system should retain the strengths of standardised examinations while incorporating structured performance tasks, analytic rubrics, individual contribution records and digital information-evaluation activities. Such a model can provide richer evidence of student competence if the instruments are valid, scoring is reliable and opportunities to demonstrate skills are equitable. In the eighth Imo State schools studied, students scored lower on each performance-based measure than on the standardised test, with digital literacy the lowest. The framework should be piloted further and evaluated with larger samples across South-Eastern Nigeria before wider claims about its effectiveness are made.

Recommendations

Based on the findings and the literature reviewed, the following recommendations are made:

I. **Combine assessment methods:** Complement conventional tests with tasks requiring evidence evaluation, collaborative problem-solving and digital information assessment.
II. **Use transparent rubrics:** Define each competency through observable criteria and share the criteria with students before assessment.
III. **Assess individual contributions:** Combine group outcomes with individual observations and reflections.
IV. **Strengthen teacher capacity:** Train teachers in rubric development, performance scoring, moderation and formative feedback.
V. **Promote digital equity:** Provide access to appropriate devices and connectivity, with accessible alternatives where necessary.
VI. **Validate the instruments:** Pilot the tasks, examine reliability and validity evidence, and assess the consistency of performance ratings.
VII. **Extend the field evidence:** Test the framework with larger samples from schools with different resource levels across South-Eastern Nigeria before making claims about its effectiveness.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: **Edinoh Kingsley:** 0009-0005-0332-2237
Nathaniel TimAhanna Ukwuoma: 0009-0005-0332-2237
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers
Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript.

References

[1] Asemota, R., Edinoh, K. & Ejike, G. A. (2026). Designing technology-enhanced learning environments for 21st-century classrooms: An empirical study. *Procedia of Social Sciences and Humanities, 5*(1), 6–13.

[2] Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M. & Rumble, M. (2012). Defining twenty-first-century skills. In P. Griffin, B. McGaw & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17–66). Springer. https://doi.org/10.1007/978-94-007-2324-5_2

[3] Care, E. & Griffin, P. (Eds.). (2017). *Assessment and teaching of 21st century skills: Research and applications*. Springer.

[4] Chai, H., Hu, T. & Wu, L. (2024). Computer-based assessment of collaborative problem solving skills: A systematic review of empirical research. *Educational Research Review, 43*, Article 100591. https://doi.org/10.1016/j.edurev.2023.100591

[5] Cheng, C. & Ho Weatherly, K. I. (2025). Navigating the double-edged sword: A meta-analysis of the effects of digital games on 21st century skills. *Educational Research Review*. https://doi.org/10.1016/j.edurev.2025.100710

[6] Edinoh, K., Nebechukwu, A. O. & Asemota, R. (2026). Rethinking assessment and feedback in the digital age: An empirical study. *American Journal of Education and Evaluation Studies, 3*(5), 138–148. https://semantjournals.org/index.php/AJEES/article/view/4097

[7] Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. American Philosophical Association.

[8] González-Mujico, F. de L., et al. (2024). Measuring students and educator digital competence beyond self-assessment: Developing and validating two rubric-based frameworks. *Education and Information Technologies, 29*, 13299–13324. https://doi.org/10.1007/s10639-023-12363-7

[9] Griffin, P. & Care, E. (Eds.). (2015). *Assessment and teaching of 21st century skills: Methods and approach*. Springer.

[10] Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.

[11] National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. The National Academies Press. https://doi.org/10.17226/13398

[12] Nyagoga, L. M., Malik, L., Lukoko, H. O. & Liu, M. (2026). Generative AI's implications on higher education students' 21st-century skills assessment: Insights from a systematic literature review. *Artificial Intelligence in Education, 2*(1), 109–126. https://doi.org/10.1108/AIIE-02-2025-0029

[13] OECD. (2021). *Beyond academic learning: First results from the survey of social and emotional skills*. OECD Publishing. https://doi.org/10.1787/92a11084-en

[14] OECD. (2023a). *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. https://doi.org/10.1787/c74f03de-en

[15] OECD. (2023b). *PISA 2022 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

[16] OECD. (2024). *PISA 2022 results (Volume III): Creative minds, creative schools*. OECD Publishing. https://doi.org/10.1787/765ee8c2-en

[17] Pan, Q., Reichert, F., Liang, Q., de la Torre, J., et al. (2025). Measuring digital literacy across ages and over time: Development and validation of a performance-based assessment. *Education and Information Technologies, 30*, 22065–22100. https://doi.org/10.1007/s10639-025-13592-8

[18] Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu*. Publications Office of the European Union. https://doi.org/10.2760/159770

[19] Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence, 11*(3), Article 54. https://doi.org/10.3390/jintelligence11030054

[20] Timmis, S., Broadfoot, P., Sutherland, R. & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal, 42*(3), 454–476. https://doi.org/10.1002/berj.3215

[21] Vlachopoulos, D. & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation, 83*, Article 101425. https://doi.org/10.1016/j.stueduc.2024.101425

[22] Voogt, J. & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st-century competences: Implications for national curriculum policies. *Journal of Curriculum Studies, 44*(3), 299–321. https://doi.org/10.1080/00220272.2012.668938

[23] Vuorikari, R., Kluzer, S. & Punie, Y. (2022). *DigComp 2.2: The digital competence framework for citizens—with new examples of knowledge, skills and attitudes*. Publications Office of the European Union. https://doi.org/10.2760/115376