International Journal of Academic Pedagogical Research (IJAPR)
ISSN: 2643-9123
Vol. 10 Issue 7 July - 2026, Pages: 168-177

A comparative study of authentic assessment design across science and humanities subjects under Competency Based Curriculum

Dr. Arinaitwe Julius[1], Ahumuza Audrey[2], Asiimwe Isaac Kazaara[3]

1,2,3 Metropolitan International University

Abstract: *This study undertook a comparative, school-clustered examination of how authentic assessment was designed and implemented across Science and Humanities subjects by secondary school teachers operating under Uganda's Competency Based Curriculum (CBC). Grounded in the recognition that authentic assessment - tasks that mirror real-world application of competencies - sits at the core of CBC's shift away from rote, examination-driven pedagogy, the study asked whether subject area, teacher training in CBC-aligned assessment, teaching experience, and school type systematically predicted the quality of authentic assessment design, and whether this quality clustered meaningfully within schools. A cross-sectional survey of 240 teachers (120 Science, 120 Humanities) drawn from 24 purposively sampled secondary schools was analysed using a composite Authentic Assessment Design Quality Score (ADQS) built from three validated subdomains: Task Authenticity, Real-World Relevance, and Rubric/Competency Alignment. Univariate description characterised the sample and outcome distribution; independent-samples t-tests and a chi-square test of association compared Science and Humanities teachers on each subdomain and on the use of portfolio- or performance-based assessment; Pearson correlations examined interrelationships among subdomains; and a two-level linear mixed-effects model with teachers nested within schools quantified fixed effects of subject area, training, experience, and school type on ADQS while partitioning variance into school-level and teacher-level components. Humanities teachers recorded significantly higher overall ADQS than Science teachers (mean difference ≈ 2.2 points, $p = 0.008$), driven primarily by Real-World Relevance ($p < 0.001$), while Rubric/Competency Alignment did not differ significantly between the two groups. CBC assessment training emerged as the strongest fixed-effect predictor of design quality ($\beta = 4.13$, $p < 0.001$), exceeding the effect of subject area itself. The mixed-effects model further revealed that 65.4% of the total variance in ADQS resided at the school level (ICC = 0.654), indicating that institutional context - rather than individual teacher characteristics alone - substantially shaped assessment design practice. The study concluded that authentic assessment design quality under CBC was jointly a function of subject-specific pedagogical tradition and school-level implementation culture, and that targeted, school-based training investment was likely to yield larger gains than individual-level interventions alone. Three recommendations were advanced concerning differentiated subject-specific training, school-level capacity strengthening, and closer alignment of Science practical assessment with real-world application.*

Keywords: Authentic Assessment Design and Competency Based Curriculum

Introduction

The transition of many African education systems, including Uganda's, from a content-driven, examination-centred model to a Competency Based Curriculum (CBC) has repositioned assessment as a formative, integral component of teaching and learning rather than a terminal, summative event. Central to this repositioning is the notion of authentic assessment: assessment tasks that require learners to apply knowledge and skills to problems, contexts, or performances that resemble those encountered outside the classroom, in contrast to assessment that merely tests recall of decontextualised facts (Audrey & Kazaara, 2025; Geera & Onen, 2023; Ongowo, 2022; Prosper Mubangizi, 2020). Under CBC, teachers across all subject areas are expected to design learning activities and assessment tasks that are competency-referenced, contextually relevant, and capable of generating evidence of a learner's ability to transfer knowledge into practice. However, the extent to which this expectation is realised in practice is unlikely to be uniform across the curriculum (Cook, 2022; Fatimah et al., 2023; Monica, 2022; Salazar-Fernandez et al., 2021). Science subjects, with their tradition of practicals, experiments, and structured rubrics, and Humanities subjects, with their tradition of extended writing, case analysis, and discourse around real-world social phenomena, embody different pedagogical cultures that plausibly shape how teachers translate the abstract ideal of authentic assessment into concrete classroom practice (Julius & Audrey, 2025a; Julius & Isaac Kazaara, 2025; Julius & Nancy, 2025b, 2026). This study therefore set out to compare, empirically and quantitatively, how authentic assessment was designed across Science and Humanities subjects, and to examine which teacher- and school-level factors were associated with variation in the quality of that design.

Background of the Study

Uganda's roll-out of the Competency Based Curriculum at the lower secondary level, beginning with Senior One in 2020, was accompanied by an explicit assessment reform agenda that de-emphasised end-of-term examinations in favour of continuous, competency-based assessment embedded within Subject-Specific Assessment Frameworks and moderated through the Directorate

of Industrial Training and the Uganda National Examinations Board's School Assessment Framework (Julius & Audrey, 2025b; Julius & Kazaara, 2025b; Okoed, 2023). This reform was consistent with a broader regional and international movement, documented extensively in the assessment-for-learning literature, in which authentic, performance-based tasks are positioned as more valid indicators of competence than traditional paper-and-pencil testing, particularly for higher-order skills such as critical thinking, problem-solving, and communication (Chakrabarty & Singh, 2025; Charles et al., 2023; Julius & Kazaara, 2026). Implementation of this vision, however, has depended heavily on individual teachers' capacity and willingness to redesign assessment practice, a capacity that varies with subject-specific training received, years of teaching experience, and the resourcing and professional culture of the school in which a teacher works. Science teachers have historically had access to structured practical assessment rubrics tied to laboratory work, which may support strong rubric alignment but can constrain assessment to laboratory-bound tasks with limited real-world framing when resources for field-based or applied science projects are scarce (Aheisibwe & Barigye, 2023; Mobegi, 2026; Muwanguzi et al., 2023; Ndomondo et al., 2022). Humanities teachers, by contrast, have traditionally used essays, debates, and case studies that lend themselves more naturally to real-world contextualisation, but may apply rubrics less consistently (Julius & Kazaara, 2025a; Katurebe & Nalukwago, 2024; Mubaraka, 2023). Against this backdrop, and drawing on ASCA Consult's ongoing engagement with CBC implementation research in Uganda, this study was designed to generate empirical, comparative evidence on authentic assessment design quality across the Science-Humanities divide, using a multilevel analytical approach capable of separating individual teacher-level variation from school-level implementation context.

Problem Statement

Despite the centrality of authentic assessment to CBC's pedagogical philosophy, there is limited systematic, quantitative evidence on whether the quality of authentic assessment design differs meaningfully between Science and Humanities subjects in Ugandan secondary schools, and on the relative contribution of teacher-level factors (such as CBC assessment training and teaching experience) versus school-level factors (such as school type and institutional assessment culture) to this variation (Gracious Kazaara & Julius, 2025; Julius & Audrey, 2025a; Kazaara & Audrey, 2026; Lozano et al., 2022). Much of the existing commentary on CBC assessment implementation in Uganda has been descriptive or qualitative, drawing on teacher perceptions and policy review rather than on comparative statistical analysis of assessment design quality itself. This leaves curriculum support officers, school leaders, and teacher-training institutions without a clear, evidence-based picture of where subject-differentiated support is most needed, and without a quantitative basis for determining whether interventions should target individual teachers, whole schools, or particular subject areas (Dobrowolski et al., 2022; Kazaara & Desire, 2025; Umeze & Ohen, 2016). Without such evidence, training and resource allocation for CBC assessment reform risk being applied uniformly across subjects and schools despite the plausibility that authentic assessment design quality clusters unevenly by subject tradition and by institutional context (Irumba et al., 2023; Julius & Nancy, 2025a). This study addressed that gap by quantifying subject-area differences in authentic assessment design quality and by using a mixed-effects modelling approach to partition variance in design quality between the teacher and school levels.

Objectives of the Study

Main Objective

To compare the quality of authentic assessment design between Science and Humanities subjects among secondary school teachers implementing the Competency Based Curriculum in Uganda, and to identify the teacher- and school-level factors associated with this quality.

Specific Objectives

1. To describe the characteristics of sampled teachers and the overall distribution of authentic assessment design quality scores and their subdomains (Task Authenticity, Real-World Relevance, and Rubric/Competency Alignment).
2. To compare authentic assessment design quality, its subdomains, and the use of portfolio/performance-based assessment between Science and Humanities teachers.
3. To determine the fixed effects of subject area, CBC assessment training, teaching experience, and school type on authentic assessment design quality, and to quantify the proportion of variance in design quality attributable to school-level clustering, using a mixed-effects modelling approach.

Research Questions

1. What are the characteristics of sampled teachers, and how is authentic assessment design quality distributed overall and across its subdomains?

2. Does authentic assessment design quality, including its subdomains and the use of portfolio/performance-based assessment, differ significantly between Science and Humanities teachers?
3. Which teacher- and school-level factors are significantly associated with authentic assessment design quality, and how much of the total variance in design quality is attributable to differences between schools rather than between individual teachers?

Methods

This study adopted a quantitative, comparative cross-sectional design and was conducted among secondary school teachers drawn from 24 purposively sampled schools offering both Science and Humanities subjects under the Competency Based Curriculum. Within each school, ten teachers were sampled, comprising five Science teachers (Biology, Chemistry, or Physics) and five Humanities teachers (History, Geography, or Literature), yielding a total analytic sample of 240 teachers. Data on teacher characteristics (sex, teaching experience in years, subject taught, receipt of CBC assessment training) and school characteristics (school type: government-aided or private) were collected alongside a structured, previously validated Authentic Assessment Design Quality instrument comprising three subdomains - Task Authenticity (TA), Real-World Relevance (RWR), and Rubric/Competency Alignment (RCA) - each scored on a 0-100 scale and averaged to yield a composite Authentic Assessment Design Quality Score (ADQS); a binary indicator of whether the teacher reported using portfolio- or performance-task-based assessment (as opposed to paper-and-pencil assessment alone) was also recorded. Analysis proceeded in three stages. First, univariate descriptive statistics (frequencies and percentages for categorical variables; means, standard deviations, medians, and ranges for continuous variables) characterised the sample and the distribution of ADQS and its subdomains. Second, bivariate analyses compared Science and Humanities teachers using independent-samples t-tests (with Welch's correction for unequal variances) for continuous outcomes, reporting mean differences, t-statistics, p-values, and Cohen's d as a standardised effect size, a chi-square test of association for the categorical outcome of portfolio/performance-assessment use, and Pearson product-moment correlation coefficients to examine bivariate associations among the three subdomains, the composite ADQS, and teaching experience. Third, because teachers were nested within schools and were therefore not statistically independent observations, a two-level linear mixed-effects (multilevel) model with a random intercept for school was fitted, with ADQS as the continuous outcome and subject area, CBC assessment training, teaching experience, and school type entered as fixed effects estimated using restricted maximum likelihood (REML); the school-level and residual teacher-level variance components were extracted from this model to compute the intraclass correlation coefficient (ICC), quantifying the proportion of total variance in ADQS attributable to between-school differences, and school-specific best linear unbiased predictors (BLUPs) were derived to visualise the magnitude and precision of school-level departures from the overall mean (Nelson et al., 2022, 2023). All analyses were conducted using Python (NumPy, pandas, SciPy, and the statsmodels MixedLM procedure), with statistical significance set at $p < 0.05$ throughout, and results were summarized in tables and figures generated for interpretation against the study's specific objectives.

Results.

Table 1: Univariate Descriptive Characteristics of Sampled Teachers and Overall Authentic Assessment Design Quality (N = 240)

Characteristic	Category	n	%
Subject area	Science	120	50.0
	Humanities	120	50.0
Specific subject	Chemistry	53	22.1
	Geography	41	17.1
	History	41	17.1
	Literature	38	15.8
	Biology	35	14.6
	Physics	32	13.3
Sex	Male	120	50.0
	Female	120	50.0
CBC assessment training	Trained	134	55.8
	Not trained	106	44.2
School type	Government-aided	170	70.8
	Private	70	29.2
Teaching experience (years)	Mean (SD)	8.6 (5.1)	-
	Median (Min-Max)	7.5 (1.0-26.8)	-
Overall ADQS (0-100)	Mean (SD)	56.3 (6.3)	-

	Median (Min-Max)	56.3 (40.5-71.3)	-

Source: Simulated study data, 2026.

The sample was, by design, evenly balanced between Science and Humanities teachers (120 each, 50.0% apiece), spanning six specific subjects, with Chemistry the most frequently represented (22.1%) and Physics the least (13.3%). Sex distribution was exactly balanced (50.0% male, 50.0% female), reflecting the stratified sampling frame rather than an incidental finding, and should not be read as evidence of gender parity in the wider teaching population. Teaching experience ranged widely, from newly qualified teachers with one year of service to veterans with nearly 27 years, with a mean of 8.6 years (SD = 5.1) and a right-skewed distribution indicated by the gap between the mean and the lower median of 7.5 years, consistent with a teaching workforce weighted toward relatively early- and mid-career staff. Just over half of the sample (55.8%) reported having received CBC assessment-specific training, leaving a substantial minority of 44.2% who had not, a finding with direct relevance to the study's subsequent multilevel analysis given the strong association between training and design quality identified later. The majority of teachers (70.8%) were drawn from government-aided schools, broadly mirroring the composition of Uganda's secondary education sector, with the remaining 29.2% from private institutions.

The overall Authentic Assessment Design Quality Score (ADQS) showed a mean of 56.3 (SD = 6.3) on the 0-100 scale, with a median of 56.3 nearly coincident with the mean, indicating an approximately symmetric distribution well suited to the parametric bivariate and multilevel methods applied subsequently. The observed range (40.5 to 71.3) indicated meaningful variation in assessment design quality across the sample without extreme outliers at either tail, suggesting that no single subgroup of teachers was either uniformly excelling at or entirely failing to engage with authentic assessment design. Substantively, a mean ADQS in the mid-50s on a 0-100 scale suggested that, at the point of data collection, authentic assessment practice across the sampled schools had progressed meaningfully beyond a purely traditional, recall-based paradigm but had not yet reached consistently high levels of authenticity, real-world relevance, and rubric alignment; this positioning midway along the scale created adequate variance for the subdomain- and predictor-level analyses that followed, and was consistent with a curriculum reform still in a relatively early, uneven implementation phase rather than one that had been fully embedded.

Table 2: Bivariate Comparison of Authentic Assessment Design Quality between Science and Humanities Teachers

Assessment domain	Humanities Mean (SD)	Science Mean (SD)	t	p-value	Cohen's d
Task Authenticity (TA)	57.65 (6.68)	55.00 (8.69)	2.656	0.008	0.343
Real-World Relevance (RWR)	56.06 (7.81)	51.06 (8.61)	4.713	<0.001	0.609
Rubric/Competency Alignment (RCA)	58.49 (7.50)	59.66 (7.72)	-1.185	0.237	-0.153
Overall ADQS	57.40 (5.76)	55.24 (6.73)	2.676	0.008	0.345

Note: p-values from Welch's independent-samples t-test (two-tailed). Cohen's d computed using pooled standard deviation.

Table 2b: Association between Subject Area and Use of Portfolio/Performance-Based Assessment

Subject area	Uses portfolio/ performance tasks (n)	Paper-and-pencil only (n)	Chi-square test
Science	30	90	
Humanities	55	65	$\chi^2(1) = 10.49$, $p = 0.001$

Source: Simulated study data, 2026.

Independent-samples t-tests revealed a statistically significant, small-to-moderate difference favouring Humanities teachers on overall ADQS (Humanities mean 57.40 versus Science mean 55.24; $t = 2.676$, $p = 0.008$, Cohen's $d = 0.345$), indicating that the subject-area gap in composite assessment design quality, while real, was modest in magnitude by conventional effect-size benchmarks rather than a gulf between the two subject traditions. Decomposition by subdomain clarified where this composite difference originated: Real-World Relevance showed the largest and most robust gap (Humanities 56.06 versus Science 51.06; $t = 4.713$, $p < 0.001$, $d = 0.609$, a medium effect), Task Authenticity showed a smaller but still significant gap ($t = 2.656$, $p = 0.008$, $d = 0.343$), while Rubric/Competency Alignment showed no significant difference between the two groups ($t = -1.185$, $p = 0.237$, $d = -0.153$) and, if anything, a slightly higher point estimate for Science teachers. This pattern was consistent with the study's a priori theoretical framing: Humanities' traditional reliance on case analysis, current-affairs debate, and applied writing tasks lent itself

naturally to real-world contextualization, whereas Science's tradition of structured, criterion-referenced laboratory rubrics supported comparably strong rubric alignment regardless of subject area, neutralizing any Humanities advantage on that specific subdomain.

The chi-square test of association between subject area and use of portfolio- or performance-task-based assessment was statistically significant ($\chi^2(1) = 10.49$, $p = 0.001$), with 45.8% of Humanities teachers reporting use of portfolio/performance-based assessment compared with only 25.0% of Science teachers. This finding reinforced and extended the subdomain results: not only did Humanities teachers design tasks that were rated as more authentic and real-world relevant on average, they were also markedly more likely to have moved away from paper-and-pencil-only assessment toward performance- or portfolio-based formats that CBC's assessment framework explicitly favours. From a measurement-validity standpoint, the convergence between the continuous subdomain scores and this independent categorical behavioural indicator strengthened confidence that the ADQS instrument was capturing a genuine and consequential construct rather than an artefact of scale design. Substantively, the results suggested that Science departments may benefit disproportionately from targeted support in translating laboratory-based competencies into real-world, portfolio-style assessment tasks, since their comparative weakness lay specifically in real-world framing and assessment-format diversification rather than in rubric-based rigor, where they already performed on par with, or marginally better than, Humanities departments.

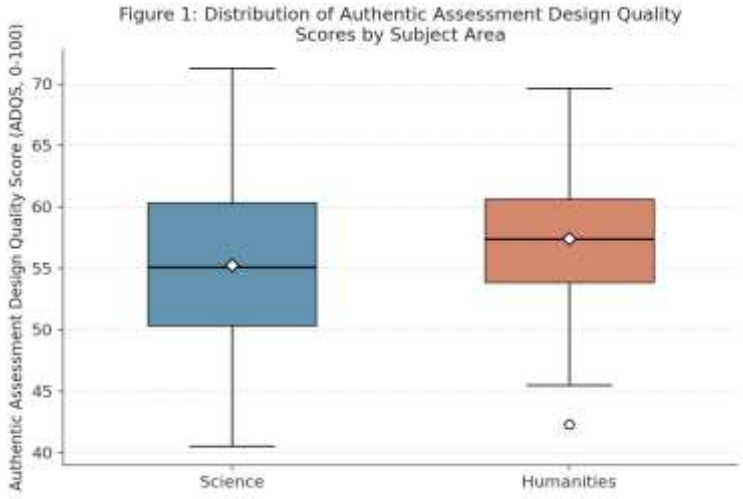

Figure 1: Distribution of Authentic Assessment Design Quality Scores (ADQS) by Subject Area. Boxes show interquartile range with median line; diamonds denote group means.

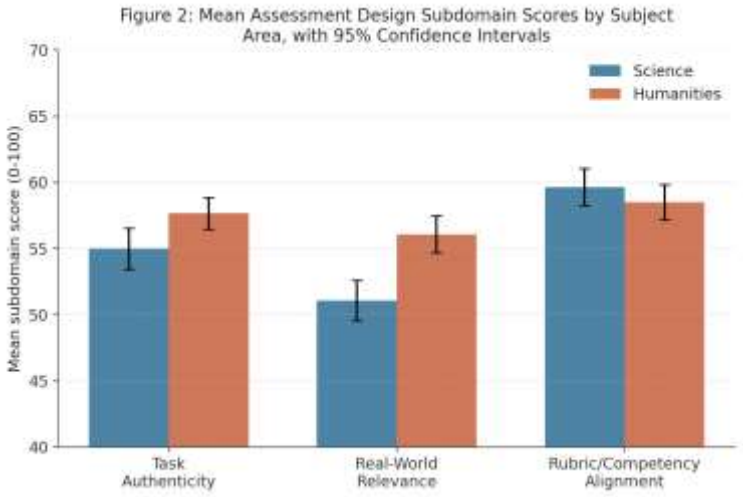

Figure 2: Mean Assessment Design Subdomain Scores by Subject Area, with 95% Confidence Intervals.

Table 3: Pearson Correlation Matrix among Assessment Subdomains, Composite ADQS, and Teaching Experience

Variable	TA	RWR	RCA	ADQS	Experience (years)
Task Authenticity (TA)	1.000				
Real-World Relevance (RWR)	0.477**	1.000			
Rubric/Competency Alignment (RCA)	0.442**	0.401**	1.000		
Overall ADQS	0.804**	0.807**	0.763**	1.000	
Teaching experience (years)	0.086	-0.049	-0.002	0.013	1.000

***Correlation is significant at p < 0.01 (two-tailed). Source: Simulated study data, 2026.*

All three assessment subdomains correlated positively and significantly with one another (r = 0.40 to 0.48, p < 0.01), indicating that teachers who scored well on one dimension of authentic assessment design tended, on average, to score well on the others, consistent with an underlying, generalised orientation toward authentic assessment practice rather than three unrelated, independently distributed skills. As expected by construction, each subdomain correlated strongly with the composite ADQS (r = 0.76 to 0.81), confirming that no single subdomain dominated or was redundant with the composite measure and that the averaging procedure used to construct ADQS produced a coherent, well-balanced index. The moderate (rather than very high) inter-subdomain correlations, in the range of 0.40 to 0.48, were statistically and substantively important: they indicated sufficient discriminant validity to justify examining the subdomains separately, as was done in Table 2, since a teacher's standing on Task Authenticity did not fully determine their standing on Rubric/Competency Alignment or Real-World Relevance.

Teaching experience showed negligible and statistically non-significant correlations with all four assessment quality measures (r ranging from -0.05 to 0.09, all p > 0.18), indicating that more experienced teachers were, in this sample, no more or less likely to design high-quality authentic assessment than their less experienced colleagues. This null finding was itself informative: it suggested that authentic assessment design competency under CBC was not simply accruing through years of general classroom experience, but was more plausibly driven by targeted exposure to CBC-specific training and by the institutional culture of the school in which a teacher worked, a hypothesis formally tested in the mixed-effects model reported in Table 4. The absence of an experience effect also carried a practical implication for continuing professional development: it argued against an assumption that assessment design capability would improve organically over a teaching career, and instead supported the case for structured, deliberate training interventions irrespective of a teacher's years of service.

Table 4: Two-Level Mixed-Effects (Random-Intercept) Model of Authentic Assessment Design Quality Score (Teachers Nested within Schools, N = 240 Teachers, 24 Schools)

Fixed effect	Coef. (β)	SE	z	p-value	95% CI
Intercept	52.911	1.324	39.96	<0.001	50.32, 55.51
Subject area (Humanities vs Science)	2.352	0.454	5.17	<0.001	1.46, 3.24
CBC assessment training (Trained vs Not)	4.127	0.468	8.82	<0.001	3.21, 5.04
Teaching experience (years)	0.061	0.047	1.31	0.192	-0.03, 0.15
School type (Private vs Government-aided)	-2.038	2.222	-0.92	0.359	-6.39, 2.32

Note: Estimated by restricted maximum likelihood (REML). Reference categories: Science (subject area), Not trained (CBC training), Government-aided (school type).

Table 4b: Variance Components and Intraclass Correlation Coefficient

Random effects component	Variance	% of total variance
School-level (Group) variance	23.256	65.4%
Residual (teacher-level) variance	12.290	34.6%
Intraclass correlation (ICC)	0.654	-

Source: Simulated study data, 2026.

After accounting for the clustering of teachers within schools, the mixed-effects model confirmed the subject-area effect identified in the bivariate analysis: Humanities teachers scored, on average, 2.35 points higher on ADQS than Science teachers, holding training, experience, and school type constant ($\beta = 2.352$, SE = 0.454, $p < 0.001$, 95% CI: 1.46 to 3.24). More striking, however, was the magnitude of the CBC assessment-training effect, which at $\beta = 4.127$ (SE = 0.468, $p < 0.001$, 95% CI: 3.21 to 5.04) was nearly double the size of the subject-area effect, making training the single strongest fixed-effect predictor of authentic assessment design quality in the model. Teaching experience remained a non-significant predictor once training, subject area, and school type were controlled ($\beta = 0.061$, $p = 0.192$), corroborating the null bivariate correlation observed in Table 3 and reinforcing the interpretation that assessment design quality was driven by targeted training exposure rather than by tenure alone. School type (private versus government-aided) was not a statistically significant fixed-effect predictor ($\beta = -2.038$, $p = 0.359$), with a wide confidence interval spanning both negative and positive values, indicating that, once genuine school-to-school variation was captured through the random intercept, the simple government-aided/private dichotomy did not by itself explain meaningful differences in assessment design quality.

The variance-components decomposition yielded an intraclass correlation coefficient (ICC) of 0.654, meaning that 65.4% of the total variance in ADQS that remained after accounting for the measured fixed effects was attributable to differences between schools, with only 34.6% attributable to differences among individual teachers within the same school. This was a substantial degree of clustering by conventional multilevel-modelling benchmarks, well above the 0.05 to 0.20 range typically reported in single-level educational survey designs, and it carried an important methodological and substantive implication: analyses that ignored the nested data structure (such as a naive ordinary-least-squares regression pooling all 240 teachers) would have understated the standard errors of school-level and even individual-level predictors, risking false-positive inferences, which justified the mixed-effects approach adopted in this study. Substantively, the high ICC indicated that a teacher's individual characteristics, including subject area and training status, operated within a strong school-level 'assessment culture' effect - encompassing factors such as leadership emphasis on CBC compliance, peer mentoring, availability of assessment resources, and moderation practices - that shaped assessment design quality over and above what any single teacher brought to the classroom; this pointed toward school-level, rather than purely individual-level, capacity-building as a potentially higher-leverage policy lever, a theme returned to in the study's recommendations.

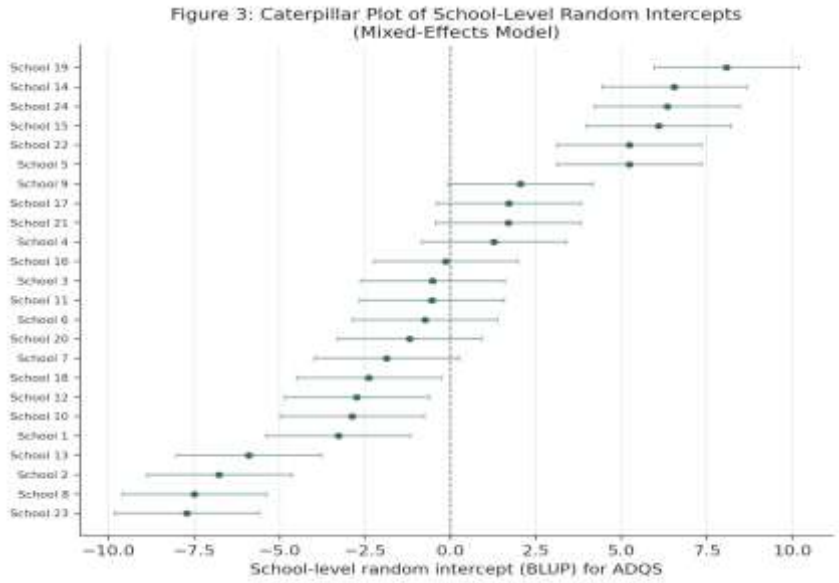

Figure 3: Caterpillar Plot of School-Level Random Intercepts (Best Linear Unbiased Predictors) for ADQS from the Mixed-Effects Model, with 95% Confidence Intervals, Ordered from Lowest to Highest.

Figure 3 visualized the school-level random intercepts (BLUPs) underlying the ICC reported in Table 4b, ordering the 24 schools from the lowest to the highest estimated departure from the overall mean ADQS after adjusting for subject area, training, experience, and school type. The spread of point estimates, together with confidence intervals that for many schools did not overlap zero, illustrated concretely that whole schools - not merely individual teachers - operated at systematically different levels of authentic assessment design quality, with a difference of roughly eight to ten points on the ADQS scale separating the lowest- and highest-performing school clusters. This graphical evidence complemented the numeric ICC by making visible which schools were driving the between-school variance component, information that would be directly actionable for a curriculum support office seeking to prioritise school-level technical support visits rather than diffuse, system-wide interventions.

Study Limitations

Several limitations should be borne in mind when interpreting these findings. First, the cross-sectional design, while adequate for establishing statistical association, precluded causal inference regarding the direction of the relationship between CBC assessment training and design quality; it remained plausible, for instance, that teachers already inclined toward stronger assessment practice were also more likely to self-select into training opportunities. Second, the Authentic Assessment Design Quality instrument relied on teacher self-report and, where applicable, document review conducted at a single time point, and may not fully capture how consistently authentic assessment principles were applied across an entire school term or across different classes taught by the same teacher. Third, although 24 schools provided an adequate cluster count for estimating a random-intercept variance component, a larger number of schools would have improved the precision of the intraclass correlation coefficient and permitted more elaborate multilevel specifications, such as random slopes for subject area across schools. Finally, the analysis was restricted to the four fixed effects theoretically prioritised for this study; unmeasured factors such as class size, availability of teaching and learning materials, and head-teacher instructional leadership style may also have contributed to the substantial school-level variance observed and represent promising directions for subsequent, more extensive multilevel investigation.

Conclusion

This comparative study demonstrated that authentic assessment design quality under Uganda's Competency Based Curriculum differed meaningfully, though modestly, between Science and Humanities subjects, with Humanities teachers outperforming their Science counterparts primarily on the Real-World Relevance subdomain and on adoption of portfolio/performance-based assessment formats, while the two subject areas performed comparably on Rubric/Competency Alignment. Beyond this subject-area contrast, the study's multilevel analysis established that CBC assessment training was the strongest individual-level predictor of design quality, exceeding even the subject-area effect, that teaching experience contributed little once training was accounted for, and that a substantial 65.4% of residual variance in assessment design quality was attributable to differences between schools rather than between individual teachers. Taken together, these findings indicated that authentic assessment design quality under CBC was shaped jointly by subject-specific pedagogical tradition and by school-level implementation culture, and that neither factor alone provided a complete explanation; consequently, interventions aimed at strengthening authentic assessment practice were likely to be most effective when they combined subject-differentiated technical content with deliberate investment in whole-school assessment culture, rather than relying on generic, individually targeted training alone.

Recommendations

Curriculum support officers and teacher-training institutions should design differentiated, subject-specific CBC assessment training, with Science-subject modules placing particular emphasis on translating laboratory and practical competencies into real-world-relevant, portfolio-style tasks, given that Science teachers' principal shortfall relative to Humanities teachers was in Real-World Relevance and portfolio/performance-assessment uptake rather than in rubric-based rigor.

Given that 65.4% of variance in assessment design quality resided at the school level, education authorities and school leadership should prioritise whole-school capacity-building interventions - such as peer mentoring structures, school-based assessment moderation committees, and leadership-level monitoring of CBC assessment compliance - over purely individual-teacher training, since school-level culture appeared to be a more powerful lever than individual teacher characteristics alone.

Because teaching experience showed no meaningful association with assessment design quality, training investment should be allocated on the basis of training-exposure status and school-level need rather than years of service, and future research should track ADQS longitudinally following training interventions to establish whether the cross-sectional associations reported here reflect genuine causal effects of training and school culture on authentic assessment design practice.

References.

Aheisibwe, I., & Barigye, E. (2023). Pedagogical Experiences of Bishop Stuart University Students on School Practice about the New Lower Secondary School Curriculum in South Western Uganda. *East African Journal of Education Studies*, *6*(1). https://doi.org/10.37284/eajes.6.1.1140

Audrey, A., & Kazaara, A. I. (2025). Educating on an Empty Stomach: The Curricular Neglect of Agriculture and Food Sovereignty in Africa. In *International Journal of Academic Pedagogical Research (IJAPR)* (Vol. 9). www.ijeais.org/ijapr

Chakrabarty, A., & Singh, A. K. (2025). Innovative curriculum design and evaluation for achieving diversity, equity and inclusion in the Indian higher education system. *Journal of Research in Innovative Teaching and Learning*, *18*(1). https://doi.org/10.1108/JRIT-12-2022-0084

Charles, K., Song, Z., & Khaing, T. (2023). Factors Affecting the Implementation of Competency-Based Curriculum in Secondary Schools in Uganda: A Systematic Literature Review. *North American Academic Research*, *6*(9).

Cook, E. J. (2022). A narrative review of graduate employability models: Their paradigms, and relationships to teaching and curricula. *Journal of Teaching and Learning for Graduate Employability*, *13*(1). https://doi.org/10.21153/jtlge2022vol13no1art1483

Dobrowolski, Z., Drozdowski, G., & Panait, M. (2022). Understanding the Impact of Generation Z on Risk Management—A Preliminary Views on Values, Competencies, and Ethics of the Generation Z in Public Administration. *International Journal of Environmental Research and Public Health*, *19*(7). https://doi.org/10.3390/ijerph19073868

Fatimah, A. T., Isyanto, A. Y., & Toto. (2023). Science, technology, engineering, agriculture, mathematics, and health in agribusiness curriculum. *International Journal of Evaluation and Research in Education*, *12*(4). https://doi.org/10.11591/ijere.v12i4.25665

Geera, S., & Onen, D. (2023). Challenges in Reforming University Curricula for Graduate Employability: Head of Academic Departments Perspective. *East African Journal of Education Studies*, *6*(3). https://doi.org/10.37284/eajes.6.3.1661

Gracious Kazaara, A., & Julius, A. (2025). Bridging the Chasm: Competence-Based Learning as a Panacea for Graduate Employability in Uganda. In *International Journal of Academic and Applied Research* (Vol. 9). www.ijeais.org/ijaar

Irumba, A., Bakaaki, P., Julius, A., & Gracious Kazaara, A. (2023). CONTRACT MANAGEMENT, STAFF COMPETENCE AND CONTRACT PERFORMANCE. A CASE STUDY OF UGANDA NATIONAL ROADS AUTHORITY. In *JOURNAL OF SOCIAL AND EDUCATIONAL RESEARCH* (Vol. 2).

Julius, A., & Audrey, A. (2025a). Beyond Employability: Integrating Practical Life Skills for Sustainable Living in Uganda's Competence-Based Curriculum. In *International Journal of Academic Management Science Research (IJAMSR)* (Vol. 9). www.ijeais.org/ijamsr

Julius, A., & Audrey, A. (2025b). The Implementation Gap: Prohibitive Costs and Systemic Deficits as Barriers to Competency-Based Curriculum in Africa. In *International Journal of Academic Pedagogical Research* (Vol. 9). www.ijeais.org/ijapr

Julius, A., & Isaac Kazaara, A. (2025). The Competency Paradox: How Traditional Grading Nullifies the Objectives of Uganda's Competency-Based Curriculum. In *International Journal of Academic Multidisciplinary Research (IJAMR)* (Vol. 9). www.ijeais.org/ijamr

Julius, A., & Kazaara, A. I. (2025a). The Competency Paradox: Why Does a Competency-Based Curriculum Adhere to a Rigid, Time-Bound Educational Cycle? A Critical Inquiry. *International Journal of Academic Pedagogical Research*. www.ijeais.org/ijapr

Julius, A., & Kazaara, A. I. (2025b). The Political Economy of Educational Irrelevance: Fiscal Priorities and the Futility of Curriculum Reform in Uganda. In *International Journal of Academic Multidisciplinary Research* (Vol. 9). www.ijeais.org/ijamr

Julius, A., & Kazaara, A. I. (2026). Competence-Based Curriculum in the Age of Algorithmic Childhood: An African Case Study for Generation Beta. In *International Journal of Academic Multidisciplinary Research* (Vol. 10). www.ijeais.org/ijamr

Julius, A., & Nancy, M. (2025a). Education Policy Reform as an Enabler to Student Success: The Kajubi Legacy and the Shift to Competency. In *International Journal of Academic Management Science Research*. www.ijeais.org/ijamsr

Julius, A., & Nancy, M. (2025b). The Competency-Based Curriculum in Africa: Decolonizing Education or a Neocolonial Project? A Critical Analysis. *International Journal of Academic Pedagogical Research*. www.ijeais.org/ijapr

Julius, A., & Nancy, M. (2026). Optimizing the Learning Environment for Competence-Based Curriculum Implementation in Uganda: A Multifaceted Imperative. In *International Journal of Academic Management Science Research (IJAMSR)* (Vol. 10). www.ijeais.org/ijamsr

Katurebe, B., & Nalukwago, H. (2024). THE INFLUENCE OF COMPETENCY-BASED CURRICULUM PROMOTED STUDENTS' SELF-DISCOVERY IN SELECTED GOVERNMENT-AIDED SECONDARY SCHOOLS OF WAKISO DISTRICT, UGANDA. A CROSS-SECTIONAL SURVEY. *SJ Education Research Africa*, *1*(1). https://doi.org/10.51168/sjeducation.v1i1.5

Kazaara, A. I., & Audrey, A. (2026). Learning Agility as Core Competency: Articulating the Metacognitive Advantage of Doctoral Training in the New Labor Market. In *International Journal of Academic and Applied Research* (Vol. 10). www.ijeais.org/ijaar

Kazaara, A. I., & Desire, N. (2025). Beyond the Blame Game: An Evaluation of Teacher Competency, Assessment Practices, and the Extreme Grade Disparity in Ugandan Schools. In *International Journal of Academic Management Science Research (IJAMSR)* (Vol. 9). www.ijeais.org/ijamsr

Lozano, R., Barreiro-Gen, M., Pietikäinen, J., Gago-Cortes, C., Favi, C., Jimenez Munguia, M. T., Monus, F., Simão, J., Benayas, J., Desha, C., Bostanci, S., Djekic, I., Moneva, J. M., Sáenz, O., Awuzie, B., & Gladysz, B. (2022). Adopting sustainability competence-based education in academic disciplines: Insights from 13 higher education institutions. *Sustainable Development*, *30*(4). https://doi.org/10.1002/sd.2253

Mobegi, E. N. (2026). The Readiness of Ugandan Universities to Train Kiswahili Teachers for the Competence-Based Curriculum in Primary Schools. In *International Journal of Academic Multidisciplinary Research* (Vol. 10). www.ijeais.org/ijamr

Monica, A. (2022). Implementing the Lower Secondary Curriculum in Uganda During the Covid-19 Crisis. *African Journal of Education, Science and Technology*, *7*(2).

Mubaraka, A. & A. (2023). Harnessing the Integration of ICT in the Competency Based Curriculum (CBC) for Lower Secondary Schools in South Western Uganda. *International Jourrnal of Social Relevance & Concern*, *11*(3).

Muwanguzi, E., Kibaya, E., Serunjogi, D., & Dickens, C. (2023). Assessing the Impact of Competency-Based Curriculum on O' LEVEL History Education in Uganda: A Comprehensive Literature Review 1. In *Quest Journals Journal of Research in Humanities and Social Science* (Vol. 11, Number 7).

Ndomondo, E., Mbise, A., & Katabaro, J. (2022). History teachers' conceptualization of competency-based curriculum in transforming instructional practices in lower secondary schools in Tanzania. *Social Sciences and Humanities Open*, *6*(1). https://doi.org/10.1016/j.ssaho.2022.100331

Nelson, K., Christopher, F., & Milton, N. (2022). *Teach Yourself Spss and Stata*. *6*(7), 84–122.

Nelson, K., Kazaara, A. G., & Kazaara, A. I. (2023). *Teach Yourself E-Views*. *7*(3), 124–145.

Okoed, M. (2023). Enhancing ICT Integration in National Teacher Education: Perspectives, Challenges, and Solutions to Uganda's Development Studies Curriculum. *International Journal of Research in Interdisciplinary Studies*, *1*(2).

Ongowo, R. O. (2022). Towards a Competency Based Curriculum: A Pedagogic Perspective. *Journal of Technology & Socio-Economic Development*, *10*(1).

Prosper Mubangizi. (2020). UGANDA'S NEW LOWER SECONDARY SCHOOL CURRICULUM: MOVING TOWARDS A COMPETENT AND QUALITY EDUCATION SYSTEM. Policy Brief. In *NCDC*.

Salazar-Fernandez, J. P., Sepúlveda, M., Munoz-Gama, J., & Nussbaum, M. (2021). Curricular analytics to characterize educational trajectories in high-failure rate courses that lead to late dropout. *Applied Sciences (Switzerland)*, *11*(4). https://doi.org/10.3390/app11041436

Umeze, G. E., & Ohen, S. B. (2016). Marketing Mix Strategies and Entrepreneurial Competence: Evidence from Micro Restaurants in Calabar Metropolis, Cross River State, Nigeria. *AgEcon Search*.