

Defining Success Criteria in Formative Assessment and Learners' Academic Performance in History and Government: Evidence from Public Secondary Schools in Kenya

Mutwota Paul Ngei¹, Dr. Charity Limboro, PhD², Dr. Ephantus Kaugi, PhD³

¹Postgraduate Student, Department of Educational Management, Policy and Curriculum Studies, Kenyatta University, Nairobi, Kenya

²Senior Lecturer, Department of Educational Management, Policy and Curriculum Studies, Kenyatta University, Nairobi, Kenya

³Lecturer, Department of Educational Management, Policy and Curriculum Studies, Kenyatta University, Nairobi, Kenya

*Corresponding Author

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000364>

Received: 06 September 2026; Accepted: 11 September 2026; Published: 27 September 2026

ABSTRACT

This article examines the influence of the definition of success criteria in formative assessment on learners' academic performance in History and Government (HAG) in Kenyan public secondary schools, taking Makueni County as the empirical case. The analysis was informed by Wiliam and Thompson's (2007) theory of formative assessment and Ramaprasad's (1983) three conditions of feedback, both of which treat a shared standard of quality as the precondition for all subsequent formative activity. A mixed-methods descriptive design was adopted. From a target population of 4,409 principals, HAG teachers and HAG students in 393 public secondary schools, Yamane's formula yielded a sample of 366, stratified across the county's nine sub-counties; 316 respondents participated, an overall return rate of 86.3%. Data were gathered using questionnaires for teachers, structured interviews for principals, 26 focus group discussions with students and a document analysis guide. It was analysed using descriptive statistics, linear regression and thematic analysis. Although a majority of teachers 113 (67.0%) strongly agreeing that they regularly define the success criteria in FA as is required, 96 (57.4%) teachers reported that the FA given to learners aligns well with stated objectives. On their side, the students differed, reporting that the success criteria are rarely articulated in advance and that assessments arrive unannounced. School-level composite means regressed on five-year KCSE averages returned $B = 2.242$, $\beta = .732$, $R^2 = 0.536$, $p < .001$, and in a four-predictor model the practice carried the smallest standardised weight of the four ($\beta = .251$, $p = .011$). The article concludes that the constraint is not the absence of success criteria but the failure to transfer them to learners, and recommends co-construction of criteria with students, exemplar-based instruction in assessment literacy, and departmental audit of criterion transfer rather than criterion articulation.

Keywords: Kenya, Makueni County, formative assessment, success criteria, History and Government, academic performance, assessment literacy, KCSE

Defining Success Criteria in Formative Assessment and Learners' Academic Performance in History and Government: Evidence from Public Secondary Schools in Kenya

By ¹Mutwota Paul Ngei, ²Dr. Charity Limboro, PhD, ³Dr. Ephantus Kaugi, PhD

INTRODUCTION

Formative assessment is the practice of continuously gathering evidence about learning while it is still in progress, and using that evidence to adjust what teachers and learners do next. Its purpose is improvement rather than certification (Cohen & Harris, 2022). Within the framework that has organized most scholarship on the subject, the first of five key strategies is the clarification, sharing and understanding of learning intentions and criteria for success (William & Thompson, 2007). The ordering is not incidental. Feedback cannot function unless the learner holds some conception of the standard being aimed at, since a comment on present work is meaningless without a reference point against which present work is judged (Ramaprasad, 1983). Success criteria supply that reference point.

History and Government make places unusual demands on this strategy. Effective achievement in HAG is demonstrated not by recall of dates and office-holders but by causal explanation, evaluation of evidence, comparison of political systems and the construction of a sustained argument, and these are qualities a learner cannot infer from the question alone. A candidate who does not know that an essay is assessed on the weighing of causes rather than the listing of them will produce a competent list, receive a low mark, and have no way of understanding why. Where the criteria of quality remain in the teacher's head, the strongest predictor of success becomes prior familiarity with academic convention.

Performance of learners in History and government is weak across a wide range of education systems. For instance, In Austria high school students attained 56.9% in national examinations against an expected 85.0% (Cotton, 2023); While in Haiti, attainment it ranged between 24.8% and 44.8% (Gedro, Hartman & Suárez, 2021); In Chile performance has remained at 36.1% despite reform efforts (Contreras & Gallardo, 2022). The pattern holds regionally: Only 33.20% of Tanzanian candidates passed in History and Civics in the Certificate of Secondary Education Examination (National Examinations Council of Tanzania, 2023), and recent Ugandan Advanced Certificate results showed a paltry 2% to 6% of candidates achieving top grades in the subject (Uganda National Examinations Board, 2025).

In Kenya, the trend is similar: The Ministry of Education (2025) reports that only a small proportion (27.4%) of candidates who sat HAG attained a grade of C+ or higher, placing it below the other social sciences. The national mean scores over the five years to 2022 remained low at 40.15% to 51.15%, corresponding to mean grades between D and C- (Kenya National Examinations Council, 2023). County data are weaker still: in Makueni County, mean scores between 2018 and 2022 ranged from 4.308 to 4.957 on the twelve-point scale, only 13% to 14.5% of candidates achieved C+ and above, and more than half scored D and below (Makueni County Director of Education Office, 2023; Ministry of Education, 2025).

Teachers in Kenya have responded by accepting formative assessment, shifting from reliance on end-of-term examinations toward questioning, short quizzes, classroom discussion, peer assessment and reflective journals (Mwanzia, 2023). The persistence of low performance alongside this adoption of FA raises a question the available literature does not answer: whether the difficulty lies in the absence of formative practice or in its quality. Success criteria are the natural place to begin, since they are the condition on which the remaining formative strategies depend.

Statement of the Problem

Teachers in Kenyan public secondary schools reportedly use various strategies to define the criteria for success during learning, and school principals are in agreement with the teachers that the definition of the success criteria is the foundational stage in FA, and that different strategies are used to realize it. At the same time, regardless of its use performance in HAG has continued to fall. The available literature does not explain the discrepancy, for three reasons. First, existing studies treat the definition of success criteria as a single undifferentiated practice. In addition, it happens in three distinct operations: communicating objectives before a task is set, preparing learners to interpret how work will be evaluated, aligning tasks with stated objectives, tracking progress against them, and applying criteria consistently. A school may perform some well and others not at all, and the aggregate measures reported by Muthoka (2024), Munguti (2024) and Onyango and Omondi (2024) cannot distinguish

such profiles. Secondly, the evidence base rests almost entirely on the reports given by the teachers. If criteria are articulated by teachers but not received by learners, teacher-reported practice registers as high while the mechanism through which criteria are supposed to operate is absent. No Kenyan study has tested this by placing teacher, principal and learner accounts of the same practice alongside one another. Third, reported effects are expressed as raw percentage improvements attributed to schools that define criteria clearly, the 5.9% reported by Onyango and Omondi (2024) in Kisumu East, the 6.8% reported by Nthenya (2023) in Machakos. Such figures estimate neither the strength of association nor the share of variation accounted for, and cannot be weighed against competing formative practices. This article addresses the three gaps by disaggregating the practice into its constituent activities, triangulating teacher, principal and learner accounts of each, and estimating the school-level association between criterion-setting and KCSE attainment in HAG.

Objective

The general objective is to examine the influence of the definition of success criteria in formative assessment on learners' academic performance in History and Government in Kenyan public secondary schools, using Makueni County as the empirical case. The specific objectives are:

1. To establish the extent to which HAG teachers undertake the constituent activities of defining success criteria during formative assessment.
2. To compare teacher, principal and learner accounts of those activities.
3. To estimate the association between school-level criterion-setting and learners' KCSE attainment in HAG.

THEORETICAL FRAMEWORK

The analysis is anchored on the theory of formative assessment advanced by Wiliam and Thompson (2007), which conceives formative assessment not as a set of classroom techniques but as a systematic process integrating assessment with instruction. Of its five key strategies, the first is the clarification, sharing and understanding of learning intentions and criteria for success. The three verbs are distinct and cumulative. Clarification is the work the teacher does alone, resolving for herself what quality looks like in the task at hand; sharing is the act of transmission; understanding is a state achieved by the learner, and it alone is treated as the operative condition. A criterion clarified and shared but not understood has not been established. This distinction supplies the analytical instrument used below: teacher reports speak to clarification and sharing, learner reports to understanding, and the two need not agree.

Ramaprasad's (1983) formulation of feedback specifies the same requirement from another direction. Feedback exists only where three conditions hold together: knowledge of the desired standard, evidence of the learner's present position, and understanding of how the gap may be closed. The first is the criterion term, and where the learner does not possess it the second and third are inert, a mark communicates position without reference, and advice on closing a gap point toward an unspecified destination. Black and Wiliam (2009) add that learners who do not hold the standard cannot regulate their own work and must depend on the teacher for every judgement of adequacy, which converts formative assessment from a developmental process into a cycle of validation. Applied to HAG, the framework predicts that where criteria are absent feedback will register as evaluative rather than instructive, and that the effect of criterion-setting on attainment will be indirect, operating through the other formative strategies rather than independently of them. Its standardised weight in a model containing those strategies should therefore be modest even where its bivariate association is strong.

Review of Related Literature

Watson (2024) argues that clear success criteria give learners an accurate reading of what is expected and thereby sustain motivation and focus. Brown (2024), surveying fourteen British secondary schools, found this to matter most in subjects such as History and Civics, where the complexity of events and institutions can overwhelm

learners: detailed criteria allowed learners to gauge their grasp of content, prioritise topics and refine the evaluative skills the subjects demand. Eagan (2023) makes the sequencing explicit, holding that comprehension depends on how learners come to terms with objectives at the outset of learning rather than at its conclusion.

The transmission problem appears repeatedly. Smith (2023) concluded that in United States high schools' assessment criteria are commonly communicated implicitly, or in language presupposing an assessment literacy learner have not been taught. Hussein (2023) found in Yemen that teachers substitute informal benchmarks and subjective judgement for specified criteria, leaving learners without a stable framework for interpreting expectations. Urdan's (2004) observational study reached the strongest version of the finding: objectives were rarely discussed adequately, and the intentions communicated were mixed and internally contradictory. Urdan further observed that learners' responses to goals varied by age and attainment level, which suggests ambiguity is not neutral in its distribution.

African evidence is consistent in direction. Opong (2024), in a phenomenological study of 113 teachers and 124 learners in Ghana, reported that clearly defined criteria enabled learners to understand what was expected and how to demonstrate it, with an associated examination improvement of 19.4%. Barrow (2024) found in Gambia that clear criteria helped learners concentrate on key concepts and develop the capacity to interpret historical events and governmental structures, while generating the diagnostic information teachers need to target support. Sotho (2025), in KwaZulu-Natal, attributed measurable improvement to criteria that allowed purposeful engagement with material, and noted a secondary benefit in the transparency of grading.

Kenyan studies concentrate on the same association. Onyango and Omondi (2024), sampling seven schools in Kisumu East Sub-County, reported a 5.9% KCSE improvement where criteria such as the ability to analyse events and articulate government policy were clearly stated, and Nthenya (2023) reported a 6.8% improvement in Machakos County. Munguti (2024), in Kitui County, added a distributional finding of some importance: irregular criteria affect lower-performing learners disproportionately, because learners with stronger self-regulatory skills can compensate for ambiguity while weaker learners cannot, so inconsistency widens existing attainment gaps. Wambua (2024) found that without explicit objectives capable learners focus on task completion rather than mastery, producing work that understates their capability. Muthoka (2022) found in Makueni County itself that learners are seldom taught how to interpret evaluative judgements, that feedback is delivered as comments or scores without explicit connection to learning goals, and that learners consequently perceive feedback as teacher-controlled rather than as a resource for improvement.

Two limitations run through this body of work. Reported effects are raw percentage differences between school groups rather than estimated associations, carrying no information about strength or explained variation. And with the partial exception of Muthoka (2022), the studies measure what teachers do rather than what learners receive, leaving the transmission failure Smith (2023) and Urdan (2004) identify unaddressed in the African literature.

METHODOLOGY

The study employed a mixed-methods approach within a descriptive research design, on the reasoning that quantitative data would establish the extent and patterning of practice while qualitative data would establish how it is experienced by those subject to it (Creswell, 2018). The design permits examination of assessment practices and learner outcomes in the unmanipulated school environment. Makueni County was selected purposively as the empirical case. It has a population of 987,653 across 8,008.9 square kilometres (Kenya National Bureau of Statistics, 2019), and its HAG attainment has remained below a C mean grade throughout the reference period, making it a case in which the practice of interest can be observed under conditions of sustained underperformance.

The target population comprised 393 principals, 892 HAG teachers and 3,124 HAG students in the county's 393 public secondary schools, totaling 4,409. Yamane's (1967) formula at a 5% margin of error yielded a sample of 366. Stratified sampling generated nine strata corresponding to the county's sub-counties, and within strata schools recording persistently low HAG KCSE results over the five years to 2025 were purposively selected.

The procedure produced 39 principals, 171 HAG teachers and 156 HAG students, the last drawn by simple random selection from a frame first stratified by gender so that selection within each category was proportionate to its size. Four instruments were used including a teacher questionnaire carrying five-point Likert items on the definition of success criteria, derived from the study objectives; structured interview guides with open-ended questions for principals; focus group discussions with learners in twenty-six groups of six; and a document analysis guide directing inspection of lesson plans and KCSE records, which provided an independent check on the frequency with which criteria were in fact recorded.

Instruments were piloted with 37 respondents comprising of 4 principals, 17 teachers and 16 learners, in four public secondary schools in Machakos Central Sub-County, none of whom took part in the main study. Content validity was established by university supervisors and specialists in curriculum studies, who reviewed and replaced items that failed to elicit the intended information. Reliability was established by test-retest, returning a Cronbach alpha coefficient of $r = 0.725$ (Kothari, 2005). Quantitative data were analysed in SPSS Version 25 using frequencies, percentages and composite means, and inferentially by linear and multiple linear regression at $\alpha = 0.05$. School-level composite means of teachers' ratings were paired with each school's average KCSE HAG mean score for the five years to 2024, giving 31 paired observations. Qualitative data were analysed thematically against the study objectives. Ethical clearance was obtained from the Kenyatta University Graduate School, a research permit from the National Commission for Science, Technology and Innovation, and authorisation from the County Director of Education, Makueni; consent was obtained in writing, with parental assent for learners under eighteen.

Presentation and Discussion of Findings

Of the 366 respondents sampled, 316 participated, an overall return rate of 86.3%, which exceeds the 75.0% threshold Creswell (2018) treats as adequate for generalisation.

Table 1: Response Rate

Respondents	Sampled	Participated	Return rate (%)
Principals	39	31	79.5
HAG teachers	171	168	98.2
HAG students	156	117	75.0
Total	366	316	86.3

Source: Field Data, Makueni County (2026).

Gender was distributed without marked imbalance, with 18 (58.1%) principals, 92 (54.8%) teachers and 71 (60.7%) learners male. Qualifications were high: 22 (70.9%) principals and 123 (73.2%) teachers held a bachelor's degree, and a further 6 (19.4%) and 31 (18.5%) respectively held postgraduate qualifications. Experience was substantial, with 16 (51.6%) principals having served more than fifteen years and 109 (64.9%) teachers between eleven and fifteen. The profile forecloses one explanation at the outset: these are not novice practitioners, so weak criterion-setting cannot be attributed to inexperience or inadequate qualification.

Table 2: Students' KCSE Performance in HAG in Makueni County (2020–2024)

Examination year	KCSE HAG mean score
2020	7.997
2021	7.841
2022	4.623
2023	4.535
2024	4.444

Source: Field Data (2026).

The series contains a marked discontinuity between 2021 and 2022, followed by continued decline. Whatever accounts for the break, the trajectory after it is downward in each successive year, and the post-2021 level is consistent with county figures independently reported for earlier years (Makueni County Director of Education Office, 2023; Kenya National Examinations Council, 2024).

Teachers' Accounts of Criterion-Setting Practice

Table 3: Teachers' Ratings of Success-Criteria Practices in Formative Assessment (n = 168)

Test item	SA %	A %	U %	D %	SD %
Learning objectives are communicated to students before an assignment is shared	51.7	11.3	5.6	5.6	25.8
Students are prepared to understand the criteria for success and how work is evaluated	53.7	14.8	5.6	11.1	14.8
Formative assessments are well aligned with the intended learning objectives	57.4	9.3	7.4	9.3	16.7
Student progress is tracked regularly against set standards	48.1	18.5	5.6	9.3	18.5
The definition of criteria for success is regular and has improved performance	67.0	11.1	3.4	5.6	13.0

Source: Field Data (2026).

The distribution is informative in a way that composite means conceal. On every item the modal response is strong agreement, and on the summary item 113 (67.0%) of teachers strongly agreed that criteria are regularly defined. Yet each item carries a substantial tail of strong disagreement, reaching 43 (25.8%) on advance communication of objectives. The ratings are bimodal rather than centrally clustered, indicating two distinct populations of schools rather than one population practising the activity moderately well. The weakest-endorsed practice is advance communication and the strongest is the summary judgement that criteria are defined regularly: teachers are, on their own returns, more confident that they define criteria than that they tell learners in time to be useful.

Principals corroborated the teacher accounts. One reported that formative assessment activities are shared with students through their teachers, who must explain the key objectives of the exercise (Principal, P1, School A). Another stated that teachers design assessments directly from lesson objectives and expected outcomes, so that learners are assessed on what has been taught and clearly understand the expectations (Principal, P4, School D). A third observed that where criteria are clearly stated and consistently applied, teachers know what to assess and learners comprehend what is desired of them (Principal, P8, School H).

Learners' Accounts of the Same Practice

Learners contradicted these accounts on every practice. On advance communication, a discussant reported:

At my school, the aims of formative assessment are never well explained to us in advance, so we do not get prepared. On many occasions, it is often done as a surprise. This has not helped improve our mastery of concepts in HAG, with the result being low performance in the subject.

On preparation to interpret evaluation, learners stated that they are rarely prepared for formative assessments and are not made aware of the criteria for success. On the consistency of criteria, they reported that criteria are rarely defined and that they are left to interpret what is expected of them during any assessment. The single practice on which accounts converged was progress tracking, where learners confirmed that their academic activities are monitored frequently and that teachers follow up on the timely completion of assignments:

There is frequent monitoring of our academic activities. Teachers always check whether we undertake the tasks given and whether we complete them in time.

The pattern of the divergence, rather than its existence, carries the analytical weight. Agreement is confined to the one practice that is externally observable and administratively recorded; divergence covers every practice whose completion depends on a state achieved inside the learner. This is precisely the distinction Wiliam and Thompson (2007) draw between sharing and understanding. Teachers report accurately on what they have done, learners on what has reached them, and the gap between the two is where the strategy fails. Criteria are being announced rather than taught, and announcement presupposes the assessment literacy Smith (2023) shows learners do not possess unless it is deliberately built.

Two consequences follow. Feedback loses its function, since a learner who does not hold the standard cannot determine from a comment whether it has been met, which returns the process to the teacher-dependence Black and Wiliam (2009) describe. And the burden falls unevenly: learners with stronger self-regulatory skills can reconstruct criteria from marks and remarks while weaker learners cannot, so ambiguity widens attainment gaps rather than depressing all learners equally (Munguti, 2024). Document analysis supported the learner account, in that lesson plans recorded objectives but rarely the criteria by which work would be judged, and there was little evidence of exemplars or rubrics having reached learners' hands.

Association with KCSE Attainment

School-level composite means of teachers' ratings were paired with each school's average KCSE HAG mean score for 2020 to 2024 across the 31 sampled schools and analysed by linear regression.

Table 4: Regression of Learners' Academic Performance in HAG on the Definition of Success Criteria (n = 31 schools)

Model	B	Std. Error	Beta	t	Sig.
(Constant)	-2.627	1.492		-1.760	.089
Definition of success criteria in FA	2.242	0.388	.732	5.783	.000

Dependent variable: Learners' academic performance in HAG

Source: Field Data (2026).

The association is positive, strong and significant. The standardised coefficient of .732 indicates that a one standard deviation increase in criterion-setting practice corresponds to roughly three-quarters of a standard deviation increase in school mean score, and $R^2 = 0.536$ indicates that 53.6% of the variation in school-level HAG attainment is accounted for. The unstandardised coefficient of 2.242 means a one-point increase on the composite scale is associated with an estimated 2.242-point increase in KCSE mean score, a substantial quantity on a twelve-point scale. The intercept of -2.627 does not differ significantly from zero ($p = .089$) and should not be interpreted substantively, a composite score of zero lying far outside the observed range.

The remaining 46.4% of variation is attributable to factors outside the model, including teacher effectiveness, learner motivation, school resources, instructional strategy, parental support and leadership practice. Two cautions attach. The estimate rests on 31 school-level observations, so it establishes covariation rather than causation, and schools whose teachers rate criterion-setting highly may differ systematically in ways the model does not capture. And the predictor is teacher self-report, which the learner data show does not correspond to learner receipt; the coefficient therefore measures the association between a school's assessment culture as its teachers understand it and its examination outcomes, rather than the effect of criteria reaching learners.

When the four formative assessment practices examined in the wider study were entered together, accounting for 87.3% of variation, the definition of success criteria carried the smallest standardised weight ($\beta = .251$, $p =$

.011), behind the provision of feedback ($\beta = .329$), classroom discussions ($\beta = .296$) and the activation of self-assessment ($\beta = .289$). The demotion is theoretically expected rather than anomalous: criterion-setting is not a mechanism acting on attainment directly but the condition that makes the other three operative, so where the practices are modelled together its work is performed through them.

CONCLUSION

The definition of success criteria is strongly associated with HAG attainment at school level, accounting for 53.6% of the variation in five-year KCSE means. On the quantitative evidence alone, the practice would appear to be a high-yield intervention. The qualitative evidence establishes that it is not currently functioning as one, and specifies why. Teacher's report performing the activity, principals confirm it, and learners deny receiving it. The denial is patterned rather than general: accounts converge on progress tracking, which is externally observable and administratively recorded, and diverge on advance communication, preparation to interpret evaluation and consistency of criteria, each complete only when a learner holds the standard. What is failing is not articulation but transfer criteria announced to learners who have not been taught to read them, the transmission problem Smith (2023) identifies and Muthoka (2022) had already documented in this county.

The theoretical framework accommodates this outcome rather than being embarrassed by it. Wiliam and Thompson (2007) make understanding, not sharing, the operative condition of the first strategy, and Ramaprasad (1983) makes the learner's possession of the standard the term without which feedback does not exist. A regime that clarifies and shares but does not secure understanding has satisfied two of three conditions and produced none of the effect. The relegation of criterion-setting to the smallest standardised weight in the four-predictor model is consistent with this: the practice operates through the other strategies rather than alongside them, so where it fails their yield is reduced without its failure being separately visible. The distributional consequence is the most serious. Learners who can reconstruct criteria from marks will do so; learners who cannot are left to infer the standard from the outcome of having missed it. Ambiguity in success criteria does not lower attainment evenly. It sorts learners by the assessment literacy they brought to the classroom, and in a subject where more than half of candidates score D and below, that is not a marginal effect.

RECOMMENDATIONS

Based on the findings and conclusions, the following recommendations are made:

- i. HAG teachers should co-construct success criteria with learners rather than announcing them. Where learners help to generate the criteria for a task, understanding is a by-product of the activity rather than something that must be separately secured afterwards. This addresses the transfer failure directly and requires no additional instructional time, since the co-construction replaces the announcement, it displaces.
- ii. Criteria should be taught through annotated exemplars rather than stated in prose, since learners in this study could not interpret criteria expressed as statements — the assessment-literacy deficit Smith (2023) describes. A pair of scripts at different standards, annotated to show which qualities earn credit, conveys a criterion that a sentence cannot. Departments should build and maintain banks of such exemplars for recurring HAG essay types.
- iii. School leadership should audit criterion transfer rather than criterion articulation. Present monitoring inspects lesson plans, which record objectives and are therefore satisfied by articulation alone. The finding that teacher and principal accounts agreed while learner accounts diverged indicates that existing monitoring cannot detect the failure. Asking a sample of learners what a forthcoming assessment requires of them would detect it in minutes.
- iv. Advance communication of objectives should be the priority among the constituent practices. It carried the weakest teacher endorsement, drew the sharpest learner contradiction, and is the practice on which the others depend, since criteria supplied after a task has been attempted cannot inform the attempt. The

Ministry of Education and county education offices should reflect this in professional development for HAG teachers.

- v. Further research should test criterion transfer directly rather than through teacher self-report, using learner-side measures such as the accuracy with which learners predict their own marks. Intervention studies comparing co-constructed with teacher-announced criteria would establish whether the association reported here reflects a causal effect, and disaggregation by prior attainment would test Munguti's (2024) finding that criterion ambiguity widens rather than uniformly depresses attainment.

REFERENCES

1. Barrow, F. (2024). The role of clear success criteria in promoting critical thinking and academic performance in Gambia. *Journal of Educational Development*, 22(3), 45–61.
2. Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Journal of Personnel Evaluation in Education*, 21(1), 5–31.
3. Brown, J. M. (2024). *Designing formative assessments for continuous learning improvement*. Routledge.
4. Cohen, A., & Harris, M. (2022). *Formative assessment: A tool for improving learning*. Educational Press.
5. Contreras, D., & Gallardo, P. (2022). Academic performance in History and Social Studies: A case study from Chile. *International Journal of Educational Research*, 40(3), 87–99.
6. Cotton, R. (2023). Students' performance in national examinations: A report on Austrian secondary schools. *Journal of Educational Performance*, 15(4), 45–58.
7. Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.
8. Eagan, T. (2023). Unlocking student success: The power of success criteria, relationships, and clarity. *NECTFL Review*, 90, 95–102.
9. Gedro, J., Hartman, S., & Suárez, A. (2021). The state of History, Geography, and Civics education in Haitian secondary schools. *Global Journal of Education Studies*, 28(1), 34–47.
10. Hussein, A. K. (2023). Teachers' use of success criteria in formative assessment and its influence on learner motivation. *International Journal of Educational Assessment*, 15(2), 45–61.
11. Kabaka, A. (2022). *The role of formative assessment in improving student outcomes in Uganda* [Unpublished doctoral dissertation]. University of Kampala.
12. Kenya National Bureau of Statistics. (2019). *2019 Kenya population and housing census*. Government Printer.
13. Kenya National Examinations Council. (2023). *KCSE examination performance report*. KNEC.
14. Kenya National Examinations Council. (2024). *2023 KCSE examination statistics*. <https://www.knec.ac.ke>
15. Kothari, C. R. (2005). *Research methodology*. New International Publishers.
16. Makueni County Director of Education Office. (2023). *KCSE performance analysis report for Makueni County (2020–2023)*. Ministry of Education.
17. Ministry of Education. (2025). *Secondary school examination performance report*. Government of Kenya.
18. Munguti, P. N. (2024). Influence of success criteria clarity on academic performance among secondary school students in Kitui County, Kenya. *Journal of Educational Research in Africa*, 9(1), 77–94.
19. Musila, R. M. (2024). Monitoring learner progress and academic achievement in Kenyan secondary schools. *African Journal of Education Studies*, 12(3), 101–118.
20. Muthoka, L. M. (2022). Student assessment literacy and feedback practices in secondary schools in Makueni County. *Journal of Educational Studies*, 6(2), 59–75.
21. Muthoka, L. M. (2024). Communication of assessment expectations and learner engagement in History and Government. *Journal of Educational Studies*, 8(1), 44–60.
22. Mwanzia, M. M. (2023). *Influence of formative assessment practices on students' academic achievement in History and Government in Kenyan secondary schools* [Unpublished master's thesis]. University of Nairobi.

23. National Examinations Council of Tanzania. (2023). Certificate of Secondary Education Examination (CSEE) performance report. NECTA.
24. Nthenya, S. K. (2023). Continuous assessment practices and instructional responsiveness in public secondary schools. *East African Journal of Education and Social Sciences*, 4(1), 33–48.
25. Onyango, P. O., & Omondi, D. O. (2024). Alignment of formative assessment with learning objectives in history and government instruction in Kisumu East Sub-County. *Journal of Curriculum and Instruction*, 8(2), 66–84.
26. Opong, R. (2024). Formative assessment and success criteria in Ghana: A phenomenological study. *Educational Research Journal*, 34(2), 123–145.
27. Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science*, 28(1), 4–13.
28. Smith, R. T. (2023). Assessment literacy and students' understanding of success criteria in U.S. high schools. *Assessment in Education: Principles, Policy & Practice*, 30(4), 389–405.
29. Sotho, P. (2025). The influence of clear success criteria on learners' understanding and performance in KwaZulu-Natal, South Africa. *Journal of Education*, 29(1), 56–73.
30. Uganda National Examinations Board. (2025). UACE and UCE national examination performance report. UNEB.
31. Urdan, T. (2004). Learner perceptions of learning objectives and success criteria in the classroom: An observational study. *Educational Psychology Review*, 16(3), 305–327.
32. Wambua, J. M. (2024). Formative assessment practices and self-regulated learning among high-achieving students in Kenyan secondary schools. *Journal of Educational Psychology in Africa*, 5(1), 21–39.
33. Watson, L. E. (2024). Clarifying learning intentions: The role of formative assessment in improving student performance. *Journal of Learning and Instruction*, 82(1), 36–54.
34. Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Lawrence Erlbaum Associates.
35. Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper & Row.