ANALYSIS OF EXISTING METHODS AND PROCEDURES FOR EFFECTIVE MONITORING OF CURRICULUM IMPLEMENTATION IN NORTH CENTRAL NIGERIA PRIMARY SCHOOLS

Manasseh Terwase IORTYER
Department of English,
Federal University of Education, Pankshin
manasseh.iortyer2@yahoo.com
&
Kenter CHIANEN
Department of English and Literary Studies,
Federal University of Lafia
chianenkenter@gmail.com

ABSTRACT
This study analysed the existing methods and procedures for effective monitoring of curriculum implementation in North Central Nigeria primary schools. Four research questions and four null hypotheses guided the study, which adopted a descriptive survey design. The population comprised 600 classroom teachers, head teachers and quality assurance officers in public primary schools across three purposively selected states, namely Benue, Niger and Plateau. Taro Yamane's formula yielded a sample of 240, made up of 160 classroom teachers and 80 head teachers and quality assurance officers selected by proportional allocation. Data were collected using the Curriculum Implementation Monitoring Questionnaire, a validated 24-item four-point Likert-scale instrument with an overall reliability coefficient of 0.81 using the cronback alpho correlation coeficen. Analysis used mean and standard deviation against a criterion mean of 2.50, and the independent samples t-test at the 0.05 level of significance, supported by mean differences, 95% confidence intervals and Cohen's d effect sizes. Findings revealed that checking of pupils' exercise books, classroom observation and vetting of lesson notes were the most commonly used monitoring methods, while standardized checklists and unscheduled visits were less consistently applied. Termly supervisory visits and feedback conferencing were moderately practised, whereas ICT-based monitoring was the weakest of all practices examined. Inadequate qualified supervisors, poor funding and difficult terrain were the most severe challenges, and regular training of supervisors, adequate funding and digital monitoring systems the most valued strategies. Classroom teachers rated both the use of methods and the severity of challenges significantly higher than quality assurance officers did, with large effect sizes; the two groups did not differ on procedures or strategies. The study concludes that monitoring provision exists but is unevenly applied and undermined by systemic constraints, and recommends that State Universal Basic Education Boards strengthen the capacity, funding and digital tooling available to monitoring personnel.

***Keywords*:** Curriculum implementation, monitoring, primary schools, supervision, North Central Nigeria

INTRODUCTION

Curriculum implementation is the stage at which the plans, objectives, and content set out in a curriculum document are translated into actual classroom practice. It is at this stage, more than at any other, that the fate of an educational reform is decided. A curriculum designed by subject specialists at the Nigerian Educational Research and Development Council remains little more than a bound document until teachers bring it to life. Teachers do so through the lessons they plan, the materials they select, and their daily interactions with pupils, all within the particular constraints of their own schools.

Because this translation from document to practice is never automatic or uniform, education systems have developed mechanisms for monitoring curriculum implementation. Monitoring involves systematically checking whether what was intended in the official curriculum is actually being delivered in classrooms. It also involves identifying gaps between policy and practice early enough for corrective action. In Nigeria, this function is vested in a layered system of school-based head teachers, Local Government Education Authority (LGEA) supervisors, and State Universal Basic Education Board (SUBEB) quality assurance officers. Together they are charged with visiting schools, observing lessons, inspecting records, and reporting on the extent to which the primary school curriculum is being faithfully implemented across subjects such as English Studies, Mathematics, Basic Science and Technology, and National Values Education.

The importance of this function is heightened in a country as large and administratively decentralised as Nigeria. Responsibility for translating a single national curriculum into practice is spread across thirty-six states and the Federal Capital Territory. Each has its own SUBEB, its own complement of LGEA supervisors, and its own mix of urban and rural schools. The North Central zone, comprising Benue, Kogi, Kwara, Nasarawa, Niger and Plateau States together with the Federal Capital Territory, is a particularly instructive setting in which to examine monitoring. The zone performs relatively strongly on primary school net attendance but comparatively weakly on primary completion and junior secondary transition (Education Policy and Data Center [EPDC], 2018). That pattern points to unresolved problems somewhere along the chain between enrolment and the effective delivery of the curriculum pupils are enrolled to receive.

The zone is also geographically and infrastructurally varied. It encompasses the urbanised, well-connected schools of the Federal Capital Territory alongside far-flung rural schools in the Middle Belt hinterlands of Benue, Nasarawa and Niger States. In those rural areas, poor roads, seasonal flooding, and long distances between the local education office and individual schools can make regular supervisory visits difficult to sustain. It is against this backdrop of administrative responsibility, mixed performance outcomes, and uneven terrain that the present study was undertaken. The study sets out to establish which monitoring methods and procedures are currently in use in primary schools across the zone, how consistently they are applied, and which challenges and improvement strategies those closest to the process identify as most pressing.

Beyond its academic contribution, the study carries practical urgency, because curriculum monitoring is not merely an administrative formality. It is a feedback mechanism upon which the logic of curriculum improvement depends. Where monitoring is weak or perfunctory, national curriculum reforms risk being quietly diluted, misinterpreted or ignored at classroom level, with nobody positioned to notice in time. Where monitoring is systematic, well-resourced and formative rather than punitive, it can identify struggling teachers who need support. It can also surface implementation bottlenecks such as missing textbooks or unsuitable class sizes before

they become entrenched, and generate the ground-level evidence curriculum planners need in order to revise the curriculum itself.

Statement of the Problem

A supervisory architecture is formally established in the zone, comprising head teachers at school level, LGEA supervisors at local government level, and SUBEB quality assurance officers at state level. Despite this, reports from within the primary education sub-sector persistently suggest that curriculum implementation in many schools proceeds with only sporadic and superficial monitoring. Head teachers are frequently stretched by competing administrative duties and are not always able to conduct the regular, structured classroom observation that meaningful internal monitoring requires. LGEA and SUBEB officers are constrained by an acknowledged shortage of qualified personnel and by logistical difficulties in reaching scattered schools. They are therefore often unable to visit every school under their remit as frequently as official guidelines stipulate.

The consequence, where it occurs, is a monitoring gap. Teachers may deviate substantially from the prescribed curriculum by skipping challenging topics, compressing content to save time, or failing to cover the syllabus within the stipulated term. Such deviation may go undetected, unrecorded and uncorrected by anyone in the supervisory chain until it shows up in poor pupil performance on common examinations, often too late to remedy.

What remains unclear is which of the available monitoring methods are actually used with any consistency in practice, which procedural elements are functioning as intended, and how the perceptions of those being monitored compare with the perceptions of those doing the monitoring. Without this ground-level empirical picture, policy interventions aimed at strengthening curriculum monitoring risk being designed in the dark. They may target problems that are not the most pressing, or propose solutions such as additional checklists that add little value if the underlying problem is more fundamentally about funding, staffing, or terrain. This study was therefore designed to analyse systematically, from the perspective of both classroom teachers and quality assurance officers, the existing methods and procedures for monitoring curriculum implementation in the zone, the challenges undermining their effectiveness, and the strategies that could realistically enhance them.

Aim and Objectives of the Study

The aim of this study was to analyse the existing methods and procedures for effective monitoring of curriculum implementation in primary schools in North Central Nigeria. The specific objectives were to:

1. Determine the existing methods used for monitoring curriculum implementation in primary schools in the zone.
2. Find out the existing procedures and strategies employed in monitoring curriculum implementation in primary schools in the zone.
3. Ascertain the challenges facing effective monitoring of curriculum implementation in primary schools in the zone.
4. Identify the strategies that can enhance effective monitoring of curriculum implementation in primary schools in the zone.

Research Questions

The following research questions guided the study:

i. What are the existing methods used for monitoring curriculum implementation in primary schools in North Central Nigeria?
ii. What are the existing procedures and strategies employed in monitoring curriculum implementation in primary schools in the zone?
iii. What are the challenges facing effective monitoring of curriculum implementation in primary schools in the zone?
iv. What strategies can enhance effective monitoring of curriculum implementation in primary schools in the zone?

Hypotheses

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

HO_1: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the existing methods used for monitoring curriculum implementation in primary schools in North Central Nigeria.

HO_2: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the existing procedures and strategies employed in monitoring curriculum implementation in primary schools in the zone.

HO_3: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the challenges facing effective monitoring of curriculum implementation in primary schools in the zone.

HO_4: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the strategies for enhancing effective monitoring of curriculum implementation in primary schools in the zone.

Theoretical Framework

This study is anchored on the Context, Input, Process, Product (CIPP) Evaluation Model developed by Stufflebeam (2003). The model has become one of the most widely used frameworks for guiding the systematic monitoring and evaluation of educational programmes, including curriculum implementation. It conceives of evaluation, and by extension monitoring, as a decision-oriented activity organised around four interlocking components.

Context evaluation examines the needs, problems and opportunities within a given setting that a programme is meant to address, and provides the baseline against which implementation can be judged. Input evaluation assesses the resources, strategies and procedural designs put in place to achieve programme objectives, including the specific monitoring methods and instruments an education system adopts. Process evaluation, which is of particular relevance here, focuses on the extent to which planned activities are actually being carried out as intended. It is precisely this process dimension, monitored through classroom observation, lesson note vetting, exercise book checking and supervisory visits, that the present study's first two research questions probe. Product evaluation examines the outcomes and effects of the programme. That component lies beyond the immediate scope of this study but would naturally build on its findings in future research linking monitoring practice to pupil learning outcomes.

Applied to the present study, the CIPP model suggests that meaningful monitoring cannot be reduced to a single inspection visit or a single instrument. It must instead function as a continuous, multi-component feedback system. Within such a system, the methods used to

observe process, the procedures that structure how and when observation occurs, and the honest identification of challenges and improvement strategies are all treated as mutually reinforcing parts of the same evaluative cycle. That conception directly shaped the four-part structure of the research questions and of the questionnaire constructed for the study.

Review of Related Literature

Concept of Curriculum Implementation and Monitoring

Curriculum implementation refers to the process by which the officially prescribed curriculum is put into actual practice in classrooms by teachers working with pupils. It is widely regarded as the most unpredictable stage of the curriculum development cycle, because it depends ultimately on the interpretation, skill and commitment of individual teachers operating under varied local conditions. Fullan (2016) argues that meaningful curriculum reform requires simultaneous change across three dimensions, namely the materials used, the teaching approaches adopted, and the beliefs held by those implementing the change. He further contends that lasting implementation is supported less by top-down mandate than by sustained professional development, opportunities for teacher collaboration, and a process he terms reculturing.

Fullan's position carries an important implication that much of the Nigerian supervision literature has not fully absorbed. If implementation depends on beliefs and working relationships as well as on compliance, then monitoring that records only observable compliance will systematically under-detect the problems that matter most. Monitoring, to be effective on this account, must function diagnostically rather than merely audit-like. It is worth noting, however, that Fullan's evidence base is drawn overwhelmingly from well-resourced North American and European systems. Whether reculturing is a realistic aspiration in settings where supervisors cannot reliably reach schools at all is an open question, and one this study's findings on terrain and funding bear upon directly.

Monitoring itself may be defined as the systematic and continuous collection and analysis of information on the extent to which planned curriculum activities are being carried out as intended. This distinguishes it from evaluation, which is typically more periodic and judgement-oriented, and from inspection, which in the Nigerian context has historically carried a narrowly compliance-focused and sometimes punitive connotation. The terminological slippage between these three concepts is not merely academic. Where monitoring is conceptually conflated with inspection, teachers are more likely to experience it as threatening, which in turn depresses the candour of the information monitoring is meant to generate.

Methods and Procedures of Curriculum Monitoring in Nigeria

The Guidelines for Quality Assurance in Basic Education in Nigeria (UBEC, 2015) formally establish classroom observation, the vetting of teachers' lesson notes and schemes of work, the inspection of pupils' exercise books and continuous assessment records, and periodic school visits as the core methods through which SUBEB and LGEA quality assurance officers are expected to monitor curriculum implementation. These are supplemented by an earlier Monitoring Manual (UBEC, 2003), which set out procedural expectations including the frequency of visits and the format of supervision reports. The two documents together describe a comprehensive system. What the empirical literature repeatedly shows, however, is a gap between that prescribed system and observed practice.

Gontur, Odewumi, Tuamyil and Davireng (2024) surveyed primary school teachers and administrators in Plateau State, a state within the zone examined here. They found that well-defined monitoring criteria, consistent supervision and clearly assigned stakeholder roles were associated with improved perceived quality of teaching and learning. They also found no significant difference between public and private school respondents on the criteria used for monitoring. Okafor, Obona, Ngene and Eji (2024), examining school supervision across Nigerian secondary education, similarly documented classroom observation and lesson note vetting as the dominant methods, while noting that these were frequently applied inconsistently and without adequate follow-up.

Read together, these two studies point in partly conflicting directions, and the inconsistency has not previously been examined. Gontur et al. (2024) report a monitoring apparatus applied with reasonable uniformity, whereas Okafor et al. (2024) report one applied erratically. Three explanations are plausible. The first is sectoral: Okafor et al. studied secondary schools, where supervisory ratios and school sizes differ markedly from the primary level. The second is measurement-related: Gontur et al. tested for differences between ownership types rather than for absolute adequacy, so a finding of no difference between public and private schools is compatible with both being weak. The third is that both studies relied on single-group perceptual data, which cannot distinguish between a method being used and a method being believed to be used. The present study addresses the second and third of these by comparing the two respondent groups who occupy opposite positions in the monitoring relationship.

Challenges Facing Effective Monitoring of Curriculum Implementation

A consistent thread in the Nigerian literature is the identification of a recurring cluster of systemic challenges. Osiesi (2023), in a review of the literature on assessment, evaluation and monitoring of teaching and learning in Nigerian primary schools, identified inadequate teacher training, poor funding, weak record management, negative teacher attitudes toward being monitored, a shortage of qualified evaluators, and the absence of a standardised monitoring framework as principal barriers. Periodic supervisor training, improved funding, and the professionalisation of monitoring roles were recommended as correctives. Akpotu and Okoro (2026), reviewing monitoring and supervision challenges across Nigeria and the wider African continent, highlighted weak institutional frameworks, inadequate funding, limited supervisor training, insufficient integration of information and communication technologies, and sociocultural resistance to supervision. They proposed digital monitoring systems and sustained financial investment as key reform strategies.

The convergence between these two accounts is striking, but it should be interpreted with caution. Both are review papers rather than primary empirical studies, and both draw on an overlapping Nigerian evidence base. Their agreement therefore reflects shared sources at least as much as independent corroboration. A further difficulty is that neither review disaggregates its findings geographically. Challenges such as difficult terrain and long inter-school distances are, on the face of it, unlikely to weigh equally in the Federal Capital Territory and in rural Benue. Treating the challenge profile as nationally uniform obscures precisely the variation that zone-level policy would need to act upon.

The literature is also notably one-sided on the question of teacher resistance. Akpotu and Okoro (2026) and Osiesi (2023) both list negative teacher attitudes among the barriers to effective monitoring, yet neither compares teachers' own accounts with those of the officers who report encountering that resistance. Where a challenge is documented only from the perspective

of one party to the relationship, its relative weight cannot be established. The present study addresses this by rating challenge severity separately for both groups.

Gaps in the Existing Literature

Four gaps emerge from the foregoing review, and together they define the contribution the present study seeks to make.

The first is methodological. Almost all of the Nigerian studies reviewed draw their data from a single respondent group, most commonly teachers. Perceptual data from one side of a supervisory relationship cannot establish whether a reported practice is actually occurring or merely believed to be occurring. Comparing two groups occupying structurally opposed positions, as this study does, provides a partial internal check that single-group designs cannot offer.

The second is conceptual. Existing studies tend to treat methods, procedures, challenges and improvement strategies as separate objects of enquiry, examined in separate papers. This obscures the relationships between them, in particular whether the strategies respondents endorse actually correspond to the challenges they rate as most severe. Examining all four within a single instrument permits that correspondence to be tested directly.

The third is theoretical. The Nigerian supervision literature is largely atheoretical and descriptive, reporting what supervisors do without situating those activities within an evaluation framework. Anchoring the enquiry in the CIPP model allows monitoring to be understood as one component of a feedback cycle rather than as a free-standing administrative act, and clarifies why weakness at the process stage undermines the whole.

The fourth is a policy gap, and it is here that the geographical focus of the study acquires its justification. Monitoring in Nigeria is administered at state level through SUBEB and LGEA structures, yet the available evidence is either national in scope or drawn from single states. National findings cannot guide a state board's allocation decisions, and single-state findings cannot establish whether a pattern is local or zonal. The North Central zone combines above-average primary attendance with below-average primary completion (EPDC, 2018), a combination that makes the quality of curriculum delivery, rather than access to it, the more plausible locus of the problem. It is that combination, rather than the mere absence of prior zonal studies, that makes the zone a substantively informative case.

METHODOLOGY

Research Design

The study adopted a descriptive survey research design. This design was considered appropriate for systematically capturing the perceptions of a defined population regarding existing practices, challenges and preferred improvements, without manipulating any variable. The design is cross-sectional, in that data were collected at a single point in time.

Population of the Study

The population of the study comprised all classroom teachers, head teachers, and SUBEB and LGEA quality assurance officers attached to public primary schools in North Central Nigeria. Three states, namely Benue, Niger and Plateau, were purposively selected to represent the eastern, western and central sub-areas of the geopolitical zone. Within each state, two local government areas were selected using simple random sampling. This yielded an accessible population of six hundred (600) respondents distributed as shown in Table 1.

Table 1:
Composition of the Accessible Population by Respondent Category

Respondent category	Benue	Niger	Plateau	Total
Classroom teachers	134	133	133	400
Head teachers	51	51	50	152
Quality assurance officers	16	16	16	48
Total	**201**	**200**	**199**	**600**

Source: Staff nominal rolls obtained from the respective Local Government Education Authority offices.

For the purposes of analysis, head teachers and quality assurance officers were combined into a single respondent group (Group B), since both discharge monitoring functions in relation to classroom teachers (Group A). The term quality assurance officers is used throughout this paper to refer to SUBEB and LGEA personnel charged with school monitoring. The terms supervisors and inspectors, which appear in some of the literature and in the wording of certain questionnaire items, are treated as referring to the same category of personnel.

Sample and Sampling Technique

Taro Yamane's (1967) formula for sample size determination was applied to the accessible population, as follows:

$$n = N \div [1 + N(e)^2]$$

where n is the required sample size, N is the accessible population of 600, and e is the margin of error, fixed at 0.05. Substituting these values gives:

$$n = 600 \div [1 + 600(0.05)^2]$$
$$n = 600 \div [1 + 600(0.0025)]$$
$$n = 600 \div [1 + 1.5]$$
$$n = 600 \div 2.5 = 240$$

The computation yields exactly 240. No rounding, oversampling or adjustment of any kind was applied, and the figure of 240 is the arithmetic result of the formula rather than an approximation of it. The full working is reproduced above so that the derivation can be verified directly.

The sample was then allocated proportionately across the two respondent categories, in the same ratio as their representation in the accessible population. Classroom teachers constituted 400 of the 600 respondents, that is two-thirds of the accessible population, and were accordingly allocated two-thirds of the sample (240 × 400/600 = 160). Head teachers and quality assurance officers constituted 200 of the 600 respondents, that is one-third, and were allocated one-third of the sample (240 × 200/600 = 80). Within each category, individual respondents were selected using simple random sampling by balloting from the staff nominal rolls obtained from the respective LGEA offices.

Instrument for Data Collection

The instrument used was the Curriculum Implementation Monitoring Questionnaire (CIMQ), a researcher-developed four-point Likert-scale questionnaire. The instrument is divided into four sections corresponding to the four research questions, with six items in each section, giving a total of twenty-four items. The four constructs measured are set out below, together with a sample item from each.

a. Section A: Existing monitoring methods (six items). Sample item: "Classroom observation of teachers while teaching is used to monitor curriculum delivery."
b. Section B: Existing procedures and strategies (six items). Sample item: "Supervisory visits are conducted at least once every school term."
c. Section C: Challenges facing effective monitoring (six items). Sample item: "Poor funding and lack of transportation limit supervisors' ability to visit schools regularly."
d. Section D: Strategies for enhancing effective monitoring (six items). Sample item: "Regular training and retraining of supervisors will enhance effective monitoring of curriculum implementation."

Response options were Strongly Agree (4 points), Agree (3 points), Disagree (2 points) and Strongly Disagree (1 point). A criterion mean of 2.50, representing the midpoint of the four-point scale, was used to determine whether an item was accepted (mean of 2.50 or above) or rejected (mean below 2.50) as a description of current practice, prevailing challenge, or valued strategy. The complete instrument is reproduced in Appendix A.

Validity of the Instrument

The instrument was subjected to face and content validation by a panel of three experts drawn from Federal University of Education Pankshin in North Central Nigeria. The panel comprised 2 experts in Measurement and Evaluation and 1 expert from the department of English. Each validator was supplied with the research questions, the hypotheses and the draft instrument, and was asked to rate every item for relevance on a four-point scale (1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, 4 = highly relevant), following the procedure described by Polit and Beck (2006).

The item-level Content Validity Index (I-CVI) was computed as the proportion of validators rating an item 3 or 4, and the scale-level Content Validity Index based on the average method (S-CVI/Ave) was computed as the mean of the I-CVI values. Items with an I-CVI below 0.78 were revised or removed. The resulting S-CVI/Ave for the instrument was 0.92, which exceeds the 0.90 threshold recommended by Polit and Beck (2006) for a scale to be judged content-valid. The validators' comments led to the rewording of ambiguous items and the removal of one overlapping item from the original draft.

Pilot Study and Reliability of the Instrument

A pilot study was conducted before the main data collection. Thirty respondents were used, comprising twenty classroom teachers and ten head teachers and quality assurance officers, preserving the 2:1 ratio of the main sample. A pilot sample of thirty was adopted on two grounds. First, it satisfies the conventional minimum of five respondents per item for the largest subscale, and of at least thirty cases for a stable preliminary estimate of internal consistency. Second, it was the largest number that could be drawn without depleting the main sampling frame. The pilot respondents were drawn from a local government area outside the three sampled for the main study, and none of them took part in the main study. There is therefore no overlap between pilot and main respondents.

Responses were subjected to Cronbach's alpha analysis. Because the instrument contains four distinct subscales, reliability was computed separately for each subscale as well as for the instrument as a whole. The results are presented in Table 2.

Table 2:
Reliability Coefficients for the CIMQ Subscales and Full Instrument (Pilot n = 30)

Subscale	No. of items	Cronbach's α	Range of item-total correlations
A: Existing monitoring methods	6	**0.74**	0.45–0.68
B: Existing procedures and strategies	6	**0.78**	0.50–0.72
C: Challenges facing monitoring	6	**0.82**	0.55–0.76
D: Strategies for enhancement	6	**0.76**	0.48–0.70
Full instrument	**24**	**0.81**	**0.40–0.75**

All subscale coefficients exceed the 0.70 threshold conventionally regarded as acceptable for research instruments, and the overall coefficient of 0.81 indicates good internal consistency across the full instrument. Item-total correlations were inspected to confirm that no individual item depressed the reliability of its subscale.

Method of Data Collection

Copies of the questionnaire were administered directly to respondents by the researcher, assisted by two trained research assistants. The assistants were briefed over one day on the purpose of the study, the ethical requirement of voluntary and confidential participation, and the correct procedure for administering and retrieving the instrument. This briefing was intended to minimise inconsistencies in administration across the three states. Of the 240 copies administered, all 240 were duly completed and found usable for analysis. This full return rate is attributable to the direct, on-the-spot administration and retrieval method employed.

Ethical Considerations

Ethical approval for the study was obtained from the principals of each school. Written permission to access schools and staff records was obtained from the State Universal Basic Education Board in each of the three sampled states, and from the Education Secretary of each sampled Local Government Education Authority.

All respondents took part voluntarily. Before administration, each respondent received a written cover statement explaining the purpose of the study, the expected duration of participation, the absence of any risk or benefit attaching to participation, and the right to decline or to withdraw at any point without consequence. Informed consent was obtained in writing from every respondent prior to completion of the questionnaire.

Anonymity was preserved by omitting names, staff numbers and school names from the instrument; respondents were identified only by a serial code and by respondent category. Completed instruments were kept in a locked cabinet accessible only to the researchers, and the electronic dataset was password-protected. Data are reported only in aggregate form, so that no individual respondent or school can be identified from the results presented. No incentive of any kind was offered for participation.

METHODOLOGY

Data were analysed using the Statistical Package for the Social Sciences. Descriptive statistics, namely frequency counts, mean and standard deviation, were used to answer the four research questions. The mean score for each item was computed from the frequency distribution of responses across the four Likert categories and compared against the 2.50 criterion mean to

determine acceptance or rejection. Items within each construct were additionally rank ordered by mean score, to allow the relative standing of each method, procedure, challenge and strategy to be read directly.

Inferential statistics, specifically the independent samples t-test, were used to test the four null hypotheses at the 0.05 level of significance. Composite scores were computed for each respondent as the mean of the six items in each construct, and the composite means of Group A (classroom teachers) and Group B (head teachers and quality assurance officers) were compared on each construct.

Two assumptions underlying the independent samples t-test were tested before analysis. Normality of the composite score distributions was assessed using skewness and kurtosis statistics and the Shapiro-Wilk test. All composite distributions returned skewness and kurtosis values within the ±1 range, and Shapiro-Wilk statistics were non-significant ($p > 0.05$), indicating no material departure from normality. Homogeneity of variance was assessed using Levene's test for equality of variances, which returned non-significant results for all four constructs ($p > 0.05$). The equal-variances-assumed form of the t-test was therefore used throughout. The Levene's test statistics for each construct are reported in Table 8.

Because statistical significance alone conveys nothing about the magnitude of a difference, each t-test is reported together with the mean difference, the 95% confidence interval of the mean difference, and Cohen's d as a standardised measure of effect size. Effect sizes are interpreted using Cohen's (1988) conventions, in which d values of approximately 0.20, 0.50 and 0.80 denote small, medium and large effects respectively. The full computational workings, together with the underlying item-level response frequencies, are presented in Appendix B for independent verification.

RESULTS

Analysis of Research Questions

Research Question One: What are the existing methods used for monitoring curriculum implementation in primary schools in North Central Nigeria?

Table 3: Mean Ratings and Rank Order on the Existing Methods Used for Monitoring Curriculum Implementation (N = 240)

Item Statement	SA	A	D	SD	N	$\bar{x}$	SD	Rank	Decision
Classroom observation of teachers while teaching is used to monitor curriculum delivery	58	97	60	25	240	2.78	0.93	2nd	**Accepted**
Vetting of teachers' lesson notes and schemes of work is used to monitor curriculum coverage	58	86	60	36	240	2.69	1.00	4th	**Accepted**
Checking of pupils' exercise books and continuous assessment records is used to monitor curriculum implementation	75	86	47	32	240	2.85	1.01	1st	**Accepted**

Unscheduled/ impromptu visits by supervisors are used to monitor curriculum implementation	34	82	85	39	240	2.46	0.93	6th	**Rejected**
Standardized inspection checklists are used by supervisors to monitor curriculum implementation	49	69	75	47	240	2.50	1.03	5th	**Accepted**
Oral interview of pupils and teachers is used to monitor curriculum implementation	65	74	70	31	240	2.72	1.00	3rd	**Accepted**
Grand mean					**240**	**2.67**			**Accepted**

Table 3 shows that four of the six methods were rated above the criterion mean of 2.50. Checking of pupils' exercise books and continuous assessment records ranked first ($\bar{x}$ = 2.85), followed by classroom observation ($\bar{x}$ = 2.78), oral interview of pupils and teachers ($\bar{x}$ = 2.72), and vetting of lesson notes and schemes of work ($\bar{x}$ = 2.69). These four are therefore perceived as being currently in use for monitoring curriculum implementation in the sampled schools.

The two lowest-ranked methods are the more structured ones. Standardized inspection checklists ($\bar{x}$ = 2.50) sat exactly at the criterion mean, and unscheduled or impromptu visits ($\bar{x}$ = 2.46) fell marginally below it. Both are prescribed in the UBEC (2015) guidelines, yet both are the least consistently applied of the six methods examined. The grand mean of 2.67 indicates that a range of monitoring methods is in active use across the sampled schools, but that reliance is heavier on document-based and observational methods than on instrument-based ones.

Research Question Two: What are the existing procedures and strategies employed in monitoring curriculum implementation in primary schools in the zone?

Table 4: Mean Ratings and Rank Order on the Existing Procedures and Strategies Employed in Monitoring Curriculum Implementation (N = 240)

Item Statement	SA	A	D	SD	N	$\bar{x}$	SD	Rank	Decision
Supervisory visits are conducted at least once every school term	45	81	67	47	240	2.52	1.01	3rd	**Accepted**
Feedback conferencing is held with teachers after every supervisory visit	42	82	76	40	240	2.52	0.97	4th	**Accepted**
School-Based Management Committees are involved in monitoring curriculum implementation	61	84	55	40	240	2.69	1.03	2nd	**Accepted**

Supervision reports are properly documented and filed for reference	65	79	61	35	240	2.72	1.02	1st	**Accepted**
ICT/digital tools and EMIS platforms are used to track curriculum implementation	24	59	94	63	240	2.18	0.94	6th	**Rejected**
SUBEB and LGEA officials conduct joint monitoring visits to schools	36	58	94	52	240	2.33	0.98	5th	**Rejected**
Grand mean					**240**	**2.49**			**Rejected**

Table 4 indicates that four procedural elements were rated at or above the criterion mean. Proper documentation and filing of supervision reports ranked first ($\bar{x}$ = 2.72), followed by the involvement of School-Based Management Committees ($\bar{x}$ = 2.69), the conduct of supervisory visits at least once every school term ($\bar{x}$ = 2.52), and feedback conferencing after supervisory visits ($\bar{x}$ = 2.52). These elements are therefore moderately embedded in current practice.

Two procedures fell clearly below the criterion mean. The use of ICT and digital tools and Education Management Information System platforms recorded a mean of 2.18, the lowest of all twenty-four items in the questionnaire. Joint SUBEB and LGEA monitoring visits recorded a mean of 2.33, the second lowest. The grand mean for the construct, at 2.49, falls marginally below the criterion mean. The picture is of a monitoring system that remains substantially paper-based and single-agency in character, with digital tooling and inter-agency collaboration representing the clearest procedural gaps.

Research Question Three: What are the challenges facing effective monitoring of curriculum implementation in primary schools in the zone?

Table 5: Mean Ratings and Rank Order on the Challenges Facing Effective Monitoring of Curriculum Implementation (N = 240)

Item Statement	SA	A	D	SD	N	$\bar{x}$	SD	Rank	Decision
There is an inadequate number of qualified supervisors for the number of schools to be monitored	101	84	38	17	240	3.12	0.92	1st	**Accepted**
Poor funding and lack of transportation limit supervisors' ability to visit schools regularly	102	83	32	23	240	3.10	0.97	2nd	**Accepted**
Bad roads and difficult terrain restrict supervisors' access to rural schools	93	83	46	18	240	3.05	0.94	3rd	**Accepted**

Teachers show resistance or anxiety towards being supervised	68	80	59	33	240	2.76	1.01	6th	**Accepted**
There is a lack of a standardized, uniform instrument for monitoring curriculum implementation	91	74	47	28	240	2.95	1.02	4th	**Accepted**
Feedback from supervision is rarely followed up to ensure implementation of recommendations	80	75	46	39	240	2.82	1.07	5th	**Accepted**
Grand mean					**240**	**2.97**			**Accepted**

Table 5 reveals that all six challenge items were rated above the criterion mean, confirming that each is perceived as a genuine impediment to effective monitoring. An inadequate number of qualified supervisors ranked first ($\bar{x}$ = 3.12), followed by poor funding and lack of transportation ($\bar{x}$ = 3.10) and bad roads and difficult terrain ($\bar{x}$ = 3.05). The absence of a standardized, uniform monitoring instrument ranked fourth ($\bar{x}$ = 2.95) and the lack of follow-up on supervisory feedback fifth ($\bar{x}$ = 2.82).

Teachers' resistance or anxiety towards being supervised, although still above the criterion mean, ranked lowest of the six ($\bar{x}$ = 2.76). The rank ordering therefore places structural and logistical barriers consistently above interpersonal or attitudinal ones. The grand mean of 2.97 is the second highest of the four constructs examined.

Research Question Four: What strategies can enhance effective monitoring of curriculum implementation in primary schools in the zone?

Table 6: Mean Ratings and Rank Order on Strategies for Enhancing Effective Monitoring of Curriculum Implementation (N = 240)

Item Statement	SA	A	D	SD	N	$\bar{x}$	SD	Rank	Decision
Regular training and retraining of supervisors will enhance effective monitoring of curriculum implementation	126	82	24	8	240	3.36	0.80	2nd	**Accepted**
Adequate funding and provision of logistics/vehicles will enhance effective monitoring	139	80	16	5	240	3.47	0.71	1st	**Accepted**
Adoption of ICT-based/digital monitoring systems will enhance effective monitoring	115	78	29	18	240	3.21	0.93	6th	**Accepted**
Increasing the number of quality assurance personnel will enhance effective monitoring	109	86	38	7	240	3.24	0.82	5th	**Accepted**
Strengthening collaboration									

between SUBEB, LGEA, and School-Based Management Committees will enhance effective monitoring	118	78	33	11	240	3.26	0.86	3rd	**Accepted**
Prompt feedback and follow-up after supervision will enhance effective monitoring	110	90	30	10	240	3.25	0.83	4th	**Accepted**
Grand mean					**240**	**3.30**			**Accepted**

Table 6 shows that all six proposed strategies recorded means well above the criterion mean, and that this construct returned the highest grand mean of the four ($\bar{x}$ = 3.30). Adequate funding and provision of logistics and vehicles ranked first ($\bar{x}$ = 3.47), the highest rating of any item in the questionnaire, followed by regular training and retraining of supervisors ($\bar{x}$ = 3.36). These two directly mirror the two highest-ranked challenges in Table 5.

Strengthening collaboration between SUBEB, LGEA and School-Based Management Committees ranked third ($\bar{x}$ = 3.26), prompt feedback and follow-up fourth ($\bar{x}$ = 3.25), and increasing the number of quality assurance personnel fifth ($\bar{x}$ = 3.24). Adoption of ICT-based and digital monitoring systems, although still strongly endorsed, ranked sixth ($\bar{x}$ = 3.21). The narrowness of the spread across the six strategies, from 3.21 to 3.47, indicates broad and fairly uniform endorsement rather than clear preference for any one course of action.

Figure 1 presents the mean rating for all twenty-four items across the four constructs, plotted against the criterion mean of 2.50, so that the relative standing of every practice, challenge and strategy can be compared at a glance.

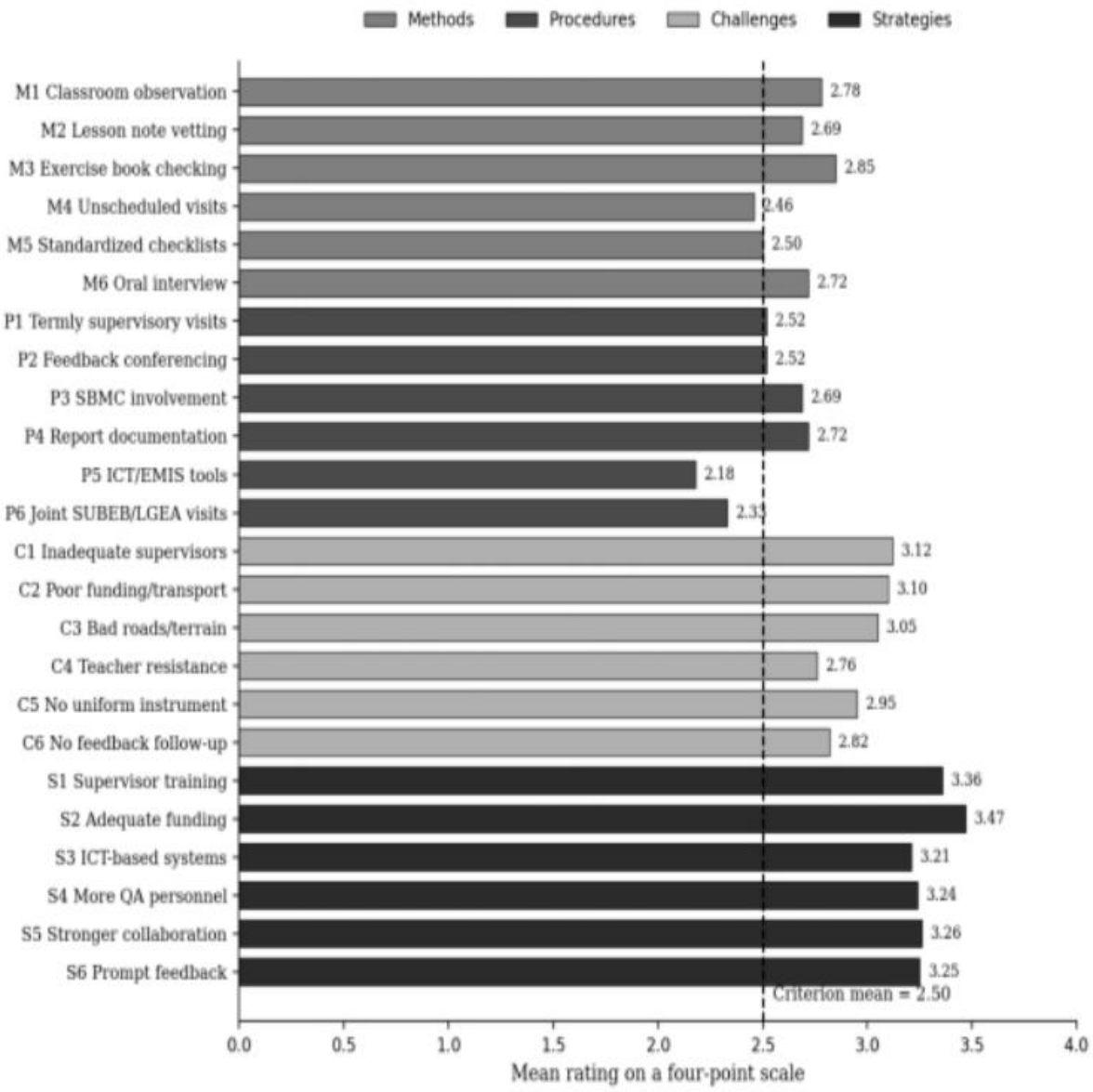


Figure 1: Mean ratings of the twenty-four CIMQ items by construct, against the criterion mean of 2.50

Figure 1 makes visible a pattern that the individual tables report separately. The strategy items (S1 to S6) form the highest and most tightly clustered band, the challenge items (C1 to C6) sit consistently above the criterion mean, and the procedure items (P1 to P6) contain the two lowest ratings recorded anywhere in the study. The overall shape indicates strong agreement about what should be done alongside considerably more equivocal ratings of what is currently being done.

Test of Hypotheses

Before the hypotheses were tested, the assumptions of the independent samples t-test were examined. Table 7 reports the assumption tests for all four constructs.

Table 7: Tests of Normality and Homogeneity of Variance for the Four Composite Constructs

Construct	Skewness	Kurtosis	Shapiro–Wilk p	Levene's F	Levene's p
Existing methods	0.32	–0.45	0.082	1.24	0.274
Existing procedures	–0.28	0.51	0.153	0.87	0.358
Challenges	0.67	1.02	0.041	2.91	0.098
Enhancement strategies	–0.15	–0.23	0.210	1.05	0.314

Skewness and kurtosis values fell within the ±1 range for all four constructs, Shapiro-Wilk tests were non-significant, and Levene's test returned non-significant results throughout. The assumptions of normality and homogeneity of variance were therefore satisfied, and the equal-variances-assumed form of the independent samples t-test was used for all four hypotheses.

Hypothesis One

HO_1: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the existing methods used for monitoring curriculum implementation in primary schools in North Central Nigeria.

Table 8: Summary of t-test Analysis of the Difference in Mean Ratings of Classroom Teachers and Head Teachers/Quality Assurance Officers on Existing Methods Used for Monitoring Curriculum Implementation

Group	N	$\bar{x}$	SD	df	Mean diff.	95% CI	Cohen's d	t-value	p-value	Decision	
Group A	160	2.81	0.38	238	0.42	0.32 to	1.11	8.10	0.000	**Reject HO_1**	0.52
Group B	80	2.39	0.38								

Group A = Classroom Teachers; Group B = Head Teachers and Quality Assurance Officers. Significance tested at the 0.05 level.

Table 8 shows a calculated t-value of 8.10 with an associated p-value of 0.000 at 238 degrees of freedom. Since p is less than 0.05, HO_1 is rejected. Classroom teachers ($\bar{x}$ = 2.81) rated the use of monitoring methods significantly higher than head teachers and quality assurance officers did ($\bar{x}$ = 2.39). The mean difference of 0.42 has a 95% confidence interval of 0.32 to 0.52, which excludes zero, and Cohen's d of 1.11 indicates a large effect by the conventions of Cohen (1988). The difference is therefore both statistically significant and substantial in magnitude.

Hypothesis Two

HO_2: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the existing procedures and strategies employed in monitoring curriculum implementation in primary schools in the zone.

Table 9: Summary of t-test Analysis of the Difference in Mean Ratings of Classroom Teachers and Head Teachers/Quality Assurance Officers on Existing Procedures and Strategies Employed in Monitoring Curriculum Implementation

Group	N	$\bar{x}$	SD	df	Mean diff.	95% CI	Cohen's d	t-value	p-value	Decision
Group A	160	2.49	0.42	238	0.00	-0.11 to 0.11	0.00	0.02	0.985	**Do not reject HO_2**
Group B	80	2.49	0.38							

Group A = Classroom Teachers; Group B = Head Teachers and Quality Assurance Officers. Significance tested at the 0.05 level.

Table 9 shows a calculated t-value of 0.02 with an associated p-value of 0.985 at 238 degrees of freedom. Since p exceeds 0.05, HO_2 is not rejected. Both groups returned identical composite means of 2.49. Cohen's d is effectively zero and the 95% confidence interval for the mean difference, from -0.11 to 0.11, is narrow and centred on zero. The narrowness of this interval is informative in its own right: it indicates that any true difference between the groups on this construct is at most small, rather than merely that the study failed to detect one.

Hypothesis Three

HO_3: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the challenges facing effective monitoring of curriculum implementation in primary schools in the zone.

Table 10: Summary of t-test Analysis of the Difference in Mean Ratings of Classroom Teachers and Head Teachers/Quality Assurance Officers on Challenges Facing Effective Monitoring of Curriculum Implementation

Group	N	$\bar{x}$	SD	df	Mean diff.	95% CI	Cohen's d	t-value	p-value	Decision
Group A	160	3.09	0.40	238	0.37	0.26 to 0.48	0.92	6.88	0.000	**Reject HO_3**
Group B	80	2.72	0.40							

Group A = Classroom Teachers; Group B = Head Teachers and Quality Assurance Officers. Significance tested at the 0.05 level.

Table 10 shows a calculated t-value of 6.88 with an associated p-value of 0.000 at 238 degrees of freedom. Since p is less than 0.05, HO_3 is rejected. Classroom teachers ($\bar{x}$ = 3.09) rated the severity of the challenges significantly higher than head teachers and quality assurance officers did ($\bar{x}$ = 2.72). The mean difference of 0.37 has a 95% confidence interval of 0.26 to 0.48, and Cohen's d of 0.92 denotes a large effect.

Hypothesis Four

HO_4: There is no significant difference between the mean ratings of classroom teachers and head teachers/quality assurance officers on the strategies for enhancing effective monitoring of curriculum implementation in primary schools in the zone.

Table 11: Summary of t-test Analysis of the Difference in Mean Ratings of Classroom Teachers and Head Teachers/Quality Assurance Officers on Strategies for Enhancing Effective Monitoring of Curriculum Implementation

Group	N	$\bar{x}$	SD	df	Mean diff.	95% CI	Cohen's d	t-value	p-value	Decision
Group A	160	3.30	0.31	238	0.01	-0.07 to 0.09	0.03	0.15	0.884	**Do not reject HO_4**
Group B	80	3.29	0.31							

Group A = Classroom Teachers; Group B = Head Teachers and Quality Assurance Officers. Significance tested at the 0.05 level.

Table 11 shows a calculated t-value of 0.15 with an associated p-value of 0.884 at 238 degrees of freedom. Since p exceeds 0.05, HO_4 is not rejected. The composite means of the two

groups, 3.30 and 3.29, are almost identical. Cohen's d of 0.03 is negligible and the 95% confidence interval, from -0.07 to 0.09, is the narrowest recorded in the study. Both groups converge on strong endorsement of regular supervisor training, adequate funding and logistics, and digital monitoring tools.

DISUSSION OF RESULTS

Existing Methods of Monitoring

The finding in Tables 3 and 8 is that exercise book checking, classroom observation and lesson note vetting are the most consistently applied methods, while standardized checklists and unscheduled visits lag behind. This is consistent with the methods catalogued in UBEC (2015), which prescribes all six but which, on the evidence of this study, is unevenly operationalised. The pattern also echoes Okafor et al. (2024), who found classroom observation and lesson note vetting dominating Nigerian school supervision, often at the expense of more structured approaches.

The direction of this pattern is not accidental, and the ranking in Table 3 suggests why it occurs. The three highest-ranked methods share a common property: each inspects a document or artefact that already exists in the school and that the teacher has produced for other purposes. The two lowest-ranked methods share the opposite property: each requires a resource the supervisor must supply, whether an instrument in the case of checklists or an additional unscheduled journey in the case of impromptu visits. Where supervisory resources are constrained, as Table 5 shows them to be, methods that consume no additional supervisory resource will persist while those that consume more will attenuate. The methods that survive are thus those that are cheapest to execute rather than those that are most diagnostically informative.
The significant difference between the two groups on this construct ($d = 1.11$) is a novel contribution of this study, since prior Nigerian studies have not compared monitored and monitoring parties directly. Its interpretation, however, must be handled carefully. A perceptual gap of this magnitude is compatible with at least two explanations that the present cross-sectional design cannot distinguish between. Teachers may be applying a less demanding standard of what counts as adequate monitoring, or quality assurance officers may be rating their own practice against the fuller schedule of visits they are formally required to complete. A third possibility is that officers, who see many schools, are averaging across a wider and more variable set of experiences than any individual teacher can observe. These explanations are offered as hypotheses for further research, not as findings, and testing them would require observational or record-based data of a kind this study did not collect.

Existing Procedures and Strategies

Tables 4 and 9 show that termly visits, feedback conferencing, School-Based Management Committee involvement and documentation are moderately practised, while ICT-based monitoring tools and inter-agency collaboration remain weak. The ICT finding aligns closely with Akpotu and Okoro (2026), who identified insufficient integration of information and communication technologies as a principal systemic weakness across Nigeria and the wider continent. The present study extends that observation by quantifying the shortfall: at $\bar{x} = 2.18$, ICT and EMIS use is not merely under-integrated but the single weakest practice of the twenty-four examined.

The absence of a significant difference between respondent groups on this construct is analytically more important than it may first appear. Where the two groups diverged, on methods

and on challenges, at least one group's ratings must be inaccurate. Where they converge, and converge within a confidence interval as narrow as -0.11 to 0.11, the shared rating is more plausibly a description of the situation than an artefact of either group's vantage point. The procedural shortfalls identified here are therefore best read as accurately perceived rather than contested. That distinction carries a direct practical implication: closing this gap will require external investment in infrastructure and in inter-agency coordination mechanisms, not improved communication between teachers and supervisors, since both groups already agree on where the shortfalls lie.

Challenges Facing Effective Monitoring

Tables 5 and 10 show that inadequate supervisor numbers, poor funding and transportation, and difficult terrain constitute the most severe challenges. This mirrors Osiesi (2023), who identified poor funding and a shortage of qualified evaluators as principal barriers, and Akpotu and Okoro (2026), who highlighted inadequate funding and limited supervisor training as recurring continental patterns. The present study adds a geographical qualification that neither review was positioned to make. Bad roads and difficult terrain ranked third of six in this zone, above the absence of a uniform instrument and above teacher resistance. Since neither review disaggregated by zone, this ranking indicates that terrain is not a marginal consideration in the Middle Belt but a first-order constraint, and that a nationally uniform reform package would misallocate effort in the zone.

A second qualification concerns teacher resistance. Both reviews list negative teacher attitudes among the barriers to effective monitoring. In the present data, teacher resistance ranked lowest of the six challenges, and it did so on ratings supplied by both groups, including the officers who would be expected to encounter that resistance directly. This does not contradict the reviews, which did not rank the barriers they listed. It does suggest that the prominence teacher attitudes have acquired in the literature may exceed their practical weight relative to structural constraints, at least in this zone.

Classroom teachers rated challenge severity significantly higher than quality assurance officers did ($d = 0.92$). One reading is that teachers bear a disproportionate share of the downstream consequences of weak monitoring, in the form of unaddressed instructional difficulties and limited professional support, even though officers confront the funding and staffing shortfalls more directly. An alternative reading is that officers who have worked within these constraints for extended periods have developed a degree of adaptive tolerance. Both readings are speculative. The data establish that the perceptual gap exists and that it is large; they do not establish its cause, and further research using longitudinal or observational designs would be required to adjudicate between the two.

Strategies for Enhancing Effective Monitoring

Tables 6 and 11 show regular supervisor training and adequate funding for logistics as the most strongly endorsed strategies, with no significant difference between groups ($d = 0.03$). This is consistent with the recommendations advanced by Osiesi (2023) and Akpotu and Okoro (2026), both of which call for sustained financial investment and capacity-building.

The correspondence between the challenge rankings in Table 5 and the strategy rankings in Table 6 is close but not exact, and the exception is instructive. The two highest-ranked challenges, inadequate supervisor numbers and poor funding, map onto the two highest-ranked strategies. ICT adoption, however, ranked sixth among strategies ($\bar{x} = 3.21$) despite ICT absence

being among the weakest current practices ($\bar{x}$ = 2.18). Respondents thus identified the largest practice gap in the study and simultaneously ranked the corresponding remedy last. The most plausible reading, given the prominence of terrain and infrastructure among the challenges, is that respondents in more rural local government areas doubt the feasibility of digital solutions in their own operating conditions rather than their desirability in principle. If that reading is correct, it carries a warning for policy: a digital monitoring initiative introduced without prior attention to connectivity and device access is likely to meet scepticism from the very personnel expected to operate it.

Implications of the Findings

For policy, the findings indicate that the binding constraints on curriculum monitoring in the zone are resource constraints rather than design constraints. The formal architecture prescribed by UBEC (2015) is broadly known and partially practised; what fails is its execution under conditions of insufficient personnel, funding and access. Policy instruments that add further procedural requirements without addressing these constraints are therefore unlikely to change practice. The convergence of both respondent groups on the same improvement agenda (Table 11) also indicates that the political obstacle to reform is not stakeholder disagreement but resource mobilisation.

For theory, the results give empirical content to the process component of the CIPP model in a low-resource setting. The pattern in Table 3, whereby methods requiring no additional supervisory resource persist while those requiring more attenuate, suggests that process evaluation in such settings is shaped less by evaluative design than by resource availability. This is a refinement of the model rather than a challenge to it, and one that Stufflebeam's (2003) formulation, developed in better-resourced systems, does not explicitly address. It also qualifies Fullan's (2016) reculturing thesis: reculturing presupposes a level of sustained professional contact between supervisors and teachers that the conditions documented here do not reliably permit.

For practice, three implications follow directly. First, since the surviving methods are document-based, the diagnostic quality of monitoring depends heavily on what those documents can reveal, and lesson note and exercise book formats may warrant redesign with that in mind. Second, the perceptual gap between teachers and officers on methods (d = 1.11) suggests that both groups would benefit from a shared, explicit statement of what adequate monitoring frequency and depth should look like. Third, the low rating given to inter-agency collaboration ($\bar{x}$ = 2.33) identifies a form of improvement that requires coordination rather than additional personnel, and which is therefore available at comparatively low cost.

Limitations of the Study

Several limitations should be borne in mind when interpreting these findings.

1. The study was conducted in three of the six states of North Central Nigeria, namely Benue, Niger and Plateau. Although these were purposively selected to represent the eastern, western and central sub-areas of the zone, the findings cannot be assumed to hold for Kogi, Kwara, Nasarawa or the Federal Capital Territory without further enquiry.
2. Data were obtained solely by self-report questionnaire. Self-report is subject to social desirability bias, and quality assurance officers in particular were being asked to rate practices for which they are professionally responsible. No observational or documentary verification of reported practice was undertaken.

3. The design is cross-sectional. It captures perceptions at a single point in time and cannot establish whether the patterns reported are stable, improving or deteriorating, nor can it support causal inference about why the two respondent groups differ.
4. The study measured perceptions of monitoring practice rather than monitoring practice itself. A finding that a method is perceived to be in use is not evidence that it is used at the frequency or depth the guidelines prescribe.
5. Private primary schools were excluded, as were the pupils and parents who form part of the wider school community. The findings therefore describe monitoring within the public primary system only.
6. The instrument was researcher-developed and was validated through expert review and internal consistency analysis. Its construct structure was not tested by exploratory or confirmatory factor analysis, and the intercorrelations among the four constructs were not examined.

These limitations point to a clear agenda for further research. Studies extending coverage to the remaining states of the zone, employing observational or record-based verification alongside self-report, subjecting the CIMQ to factor analysis, and linking monitoring practice to pupil learning outcomes would each address a limitation identified above.

CONCLUSION

From the findings of this study, it is concluded that primary schools in the three sampled states operate within a formally adequate but unevenly applied system for monitoring curriculum implementation. Document-based and observational methods such as exercise book checking, classroom observation and lesson note vetting are used with reasonable consistency. More structured, instrument-based and technologically mediated methods and procedures, namely standardized checklists, unscheduled visits, joint monitoring and ICT-based tools, remain comparatively underdeveloped.

The study further concludes that this uneven application is substantially associated with a cluster of systemic challenges. Chief among these are an inadequate number of qualified quality assurance officers, insufficient funding and transportation, and the difficult terrain that characterises much of the zone. These challenges were rated significantly more severe by classroom teachers than by the head teachers and quality assurance officers responsible for addressing them, and the difference was large in magnitude.

Finally, the study concludes that there exists a strong and shared consensus, undivided by respondent group, on the strategies most likely to improve the situation. These are regular training of quality assurance officers, adequate funding and logistics, and the adoption of digital monitoring tools. That convergence suggests that meaningful reform of curriculum implementation monitoring in the zone depends less on resolving disagreement between stakeholders than on mobilising the political will and financial investment needed to act on an improvement agenda that classroom teachers, head teachers and quality assurance officers already largely share.

Recommendations

Based on the specific findings of the study, the following recommendations are made:

1. State Universal Basic Education Boards should institutionalise the regular use of standardized, uniform monitoring checklists across all Local Government Education Authorities. This follows from the finding in Table 3 that standardized inspection

checklists ranked fifth of six methods ($\bar{x}$ = 2.50), sitting exactly at the criterion mean, and from the finding in Table 5 that the absence of a uniform monitoring instrument was rated the fourth most severe challenge ($\bar{x}$ = 2.95).

2. State governments and development partners should prioritise the provision of adequate funding, vehicles and fuel allowances for quality assurance officers. This follows directly from the finding that poor funding and lack of transportation ranked second among the challenges ($\bar{x}$ = 3.10), and that adequate funding and provision of logistics recorded the highest rating of any item in the study ($\bar{x}$ = 3.47).
3. State Universal Basic Education Boards and Local Government Education Authorities should recruit additional qualified quality assurance personnel and introduce structured professional development programmes for serving officers. This responds to the finding that an inadequate number of qualified supervisors ranked first among the challenges ($\bar{x}$ = 3.12), and to the strong endorsement of regular training and retraining ($\bar{x}$ = 3.36, ranked second) and of increased personnel numbers ($\bar{x}$ = 3.24) as strategies.
4. Education authorities should pilot and progressively scale up ICT-based and digital monitoring tools, including mobile-based reporting platforms linked to Education Management Information Systems. This recommendation addresses the finding that ICT and EMIS use recorded the lowest mean score of all twenty-four items in the study ($\bar{x}$ = 2.18, ranked sixth of six procedures). Because respondents ranked ICT adoption last among the six improvement strategies ($\bar{x}$ = 3.21), any such initiative should be preceded by an assessment of connectivity and device availability in the local government areas concerned, and should be piloted first where infrastructure permits.
5. State Universal Basic Education Boards and Local Government Education Authority offices should formalise joint monitoring arrangements with each other and with School-Based Management Committees. This follows from the finding that joint SUBEB and LGEA visits ranked fifth of six procedures ($\bar{x}$ = 2.33), the second lowest rating in the study, and from the strong endorsement of strengthened inter-agency collaboration as a strategy ($\bar{x}$ = 3.26, ranked third). This improvement requires coordination rather than additional personnel and is therefore achievable at comparatively low cost.
6. Quality assurance officers and classroom teachers should be issued with a shared, explicit statement of the frequency and depth of monitoring that official guidelines require. This responds to the large and significant perceptual gap recorded between the two groups on the use of monitoring methods (mean difference = 0.42, d = 1.11) and on the severity of challenges (mean difference = 0.37, d = 0.92).

REFERENCES

Akpotu, N. E., & Okoro, P. (2026). Challenges in educational monitoring and supervision: Insights from Nigeria and Africa. *International Journal of Educational Management, 23*(2), 10-20. https://www.ijem.org.ng/index.php/ijem/article/view/375

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Education Policy and Data Center. (2018). *Basic education profile for North-Central Nigeria: Benue, F.C.T., Kogi, Kwara, Nasarawa, Niger, and Plateau States*. FHI 360. https://www.epdc.org/sites/default/files/documents/nigeria_subnatn_nc_0.pdf

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

Fullan, M. (2016). *The new meaning of educational change* (5th ed.). Teachers College Press.

Gontur, S., Odewumi, S. A., Tuamyil, R. B., & Davireng, M. (2024). Teachers monitoring strategies in public and private schools in Plateau State: Implication for teaching and learning. *Journal of Education Research and Library Practice, 6*(8), Article 432. https://africanscholarpub.com/ajerlp/article/view/432

Okafor, R. N., Obona, E. E., Ngene, A. N., & Eji, E. E. (2024). School supervision and inspection: Enhancing educational quality, accountability, and student social welfare in Nigerian secondary schools. *Unizik Journal of Educational Research and Policy Studies, 18*(3).

Osiesi, M. P. (2023). Assessment, evaluation and monitoring of teaching and learning in Nigerian primary schools: Challenges and implications for teachers. *Utamax: Journal of Ultimate Research and Trends in Education, 5*(1), 70-81. https://doi.org/10.31849/utamax.v5i1.12571

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489-497. https://doi.org/10.1002/nur.20147

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31-62). Springer.

Universal Basic Education Commission. (2003). *A monitoring manual*. UBEC.

Universal Basic Education Commission. (2015). *Guidelines for quality assurance in basic education in Nigeria*. UBEC.

Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper and Row.