



# Digital Transformation in WASSCE: A Comparative Study of Paper-Based and Computer-Based Assessment Approaches (2024 PC2)



PRESENTED

BY

**ROSEMARY OJO-ODIDE**

**SENIOR DEPUTY  
REGISTRAR  
WAEC LAGOS**



&

**ADEBAYO FANIMOKUN**

**SENIOR ASST.**

**REGISTRAR**

**WAEC LAGOS**



# Introduction

Educational assessment is crucial for evaluating students, shaping policies, and ensuring credibility of certifications. In West Africa, the West African Senior School Certificate Examination (WASSCE), a high-stakes examination by WAEC assesses the knowledge and skills of Senior Secondary students' and Private candidates across English-speaking West African countries.



# WASSCE: Migration to CBT

Traditionally, WASSCE was conducted as paper-based test (PBT). However, with the global shift towards Computer Based Tests (CBTs) and issues such as malpractice, identified human inefficiencies and risks including the possibility of physical documents being lost or damaged, there was the need to embrace the change.

WAEC Nigeria, in 2024, introduced a hybrid CBT mode of assessment known as Computer-Based West African Senior School Certificate Examination (CB-WASSCE) in its First Series of examinations for Private Candidates.





# WAEC: CB-WASSCE MODES



## HYBRID MODE

In the WASSCE for Private Candidates, 2024 - First series, objective tests were fully computer-based, while the essay and practical components adopted a hybrid approach, where questions were delivered on-screen, but candidates provided answers in traditional answer booklets.



## DUAL MODE

Candidates were given the choice of writing the examination in either the PBT or CBT mode in the WASSCE for Private Candidates, 2024 - Second series.



# Comparability of Test Mode

With the use of the dual-mode approach in the WASSCE for Private Candidates, 2024 - Second series, there was the need to perform a comprehensive evaluation to understand its impact on candidates' performance across testing modes, the susceptibility to malpractice, and the psychometric properties such as validity and reliability of the two modes of assessment.

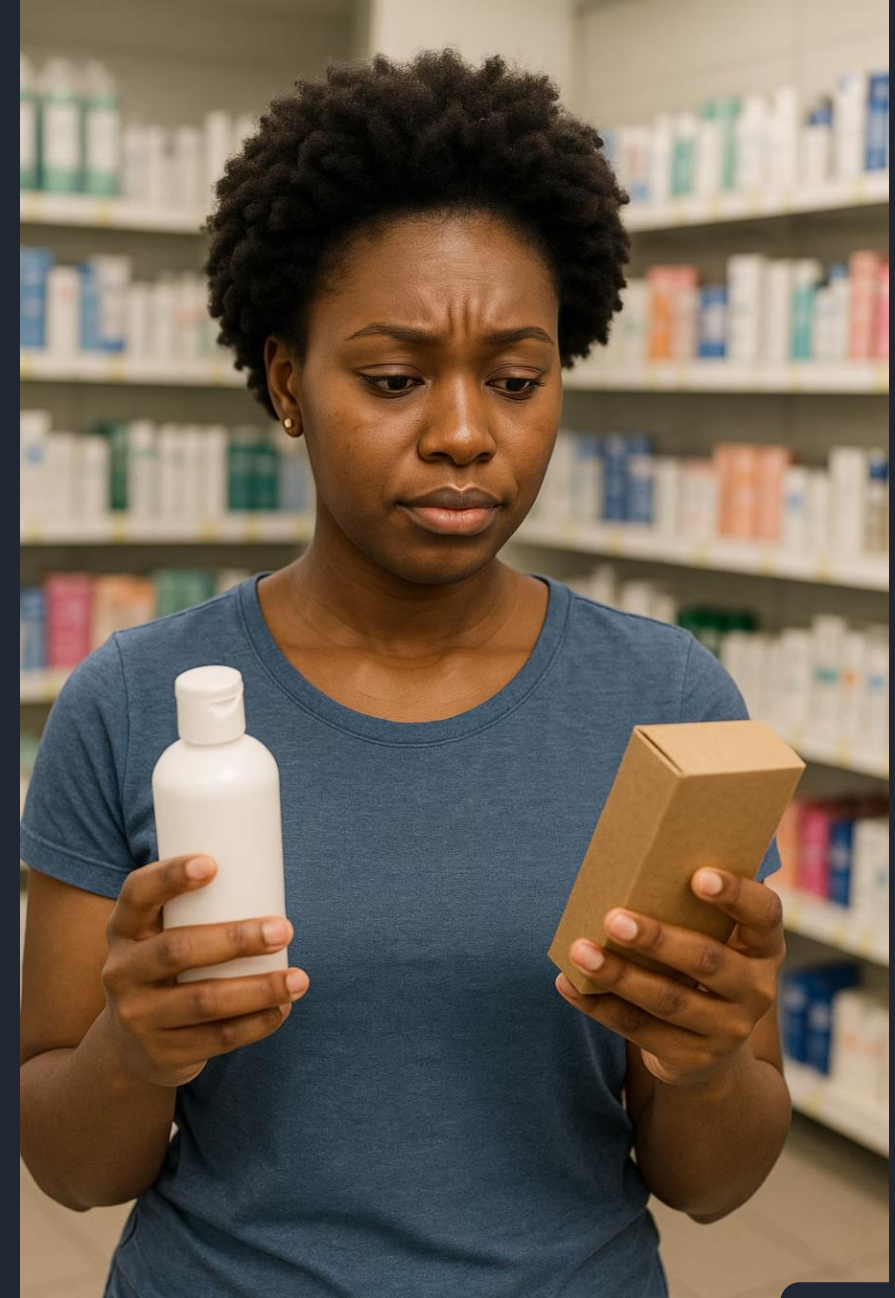
# Various Schools of Thought (PBT vs CBT)

There's no consensus in the various studies comparing PBT and CBT Test modes

In the studies conducted by Yao (2019) result revealed that test-takers performed better on the listening and writing sections on CBT compared to PBT, however there was no statistically significant difference between CBT and PBT for the reading section and total test scores.

Findings from Christopher et al. (2023) revealed a slightly higher mean achievement scores in PBT than in CBT and were significantly different among students. The reason given was that students were already used to PBT unlike CBT which was newly introduced.

Findings from Öz and Özturan (2018) indicated that there was no significant difference in the test scores between the two groups who used PBT and CBT respectively. By implication, both CBT and PBT had high reliability coefficients and internal consistency.

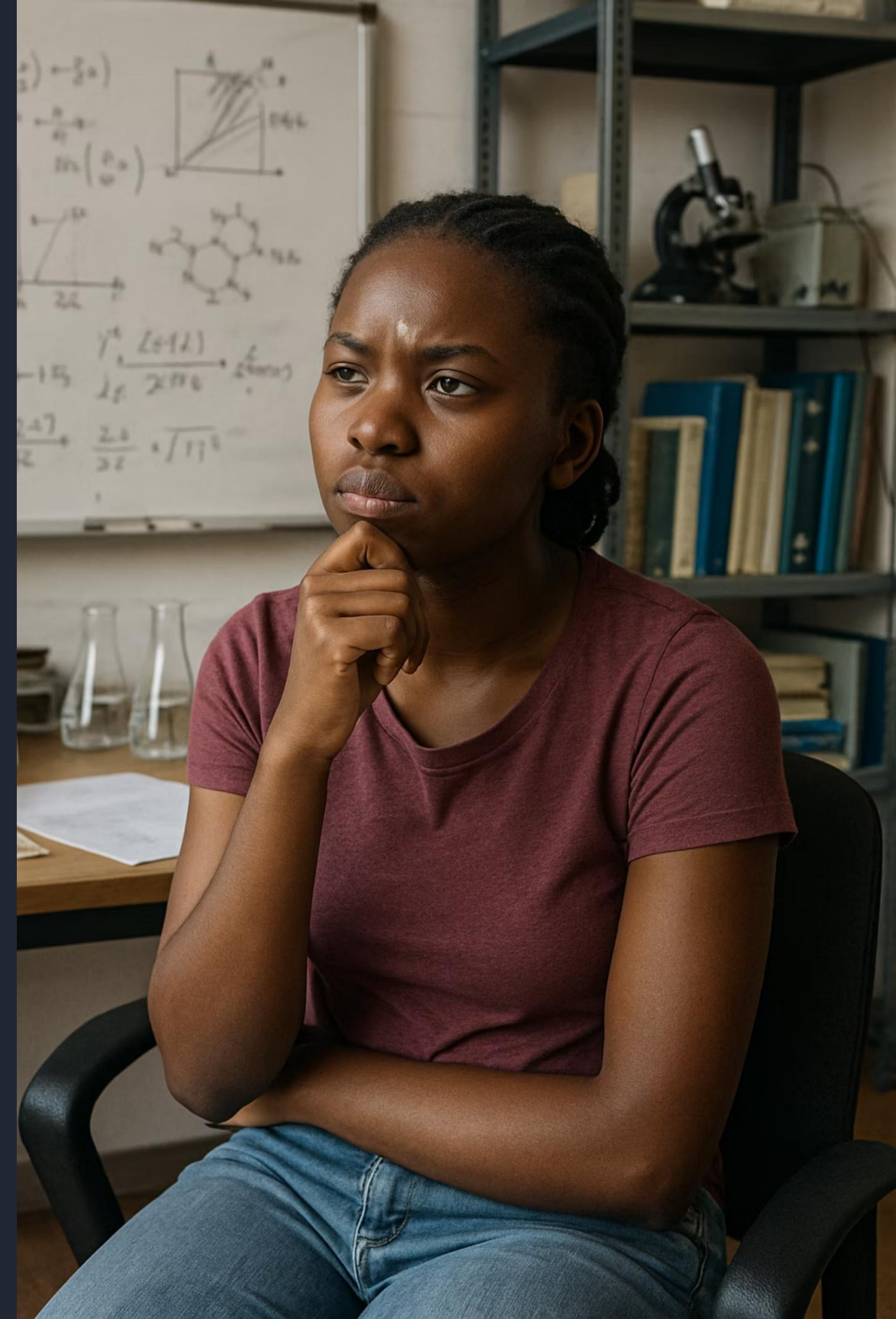




# Statement of the Problem

With the introduction of Computer-Based Testing (CBT) in WAEC examinations and the implementation of the Dual Mode in the WASSCE (PC 2) 2024, it became imperative to ensure the integrity of the examination is maintained and no undue advantage is gained by any candidate irrespective of the mode of examination taken.

To evaluate whether both testing formats ensure fairness and equivalence in assessment outcomes, the impact of examination mode (PBT or CBT) on candidates' performance, vulnerability to malpractice and test validity and reliability were examined.



# Methodology

(1) Research Design

(2) Population

(3) Sampling Technique

(4) Instruments and Instrumentation

(5) Data Analysis Procedures

Population of Candidates who sat PBT and CBT in WASSCE (PC2) 2024			
Coverage	Total Number of Candidates (PBT)		
	Male	Female	Total
NIGERIA	30288	32524	62812
%	48.22	51.77	100
Coverage	Total Number of Candidates (CBT)		
	Male	Female	Total
NIGERIA	1234	1003	2237
%	55.16	44.83	100

Ex post facto research design; 62812 PBT, 2237 CBT; Census sampling approach; Proforma designed, frequencies/percentages/t-test/mean score



# The Study

## Research Questions that guided the study

Is there a significant difference in performance rates between candidates who sat PBT and CBT in WASSCE(PC2) 2024?

How does the examination mode influence the rate of examination malpractice?

What impact does the examination mode have on the psychometric properties of test items?

How does the examination mode affect the validity and reliability of the test?



# Results and Discussion

Research question 1: Is there a significant difference in performance rates between candidates who sat PBT and CBT in WASSCE(PC2) 2024?

Indicator	Result
Performance test using 5 credit pass and above including English and Mathematics	PBT > CBT
Performance in English Language and General Mathematics	PBT > CBT
Mean score test for English Language and General Mathematics	PBT > CBT



# Performance: Candidates with 5 credit pass and above

It was observed that out of 62,812 candidates that participated in the PBT mode 33,021 (52.57%) candidates made credit score in five subjects and above including English Language and General Mathematics compared to 354 (15.82%) out of the 2,237 candidates that sat the examination using the CBT mode.

	PBT				CBT			
	Total Number Sat		5 credits and above including Gen. Mathematics and English Lang.		Total Number Sat		5 credits and above including Gen. Mathematics and English Lang	
	N	%	N	%	N	%	N	%
<b>Male</b>	30288	48.22	15839	52.29	1234	55.16	199	16.12
<b>Female</b>	32524	51.77	17182	52.82	1003	44.83	155	15.45
<b>Total</b>	62812	100	33021	52.57	2237	100	354	15.82

Source: ICTD, WAEC

# Performance in English Language and General Mathematics

Out of the 60,404 candidates that wrote English Language as PBT, 45,642 (75.56%) made C6 and above compared to 970 (46.03%) candidates out of the 2107 candidates that participated using the CBT mode. The observation was similar for General Mathematics as 40,730 (66.72%) candidates out of the 61,043 that wrote PBT made C6 and above whereas 559 (27.16%) candidates out of 2,058 candidates made C6 and above from the CBT mode.

Candidates' performance in English Language and General Mathematics in PBT and CBT modes in WASSCE(PC2) 2024 in Nigeria								
PBT								
	Total Number Sat		A1-C6		D7-E8		F9	
	N	%	N	%	N	%	N	%
English Lang.	60404	91.59	45642	75.56	8720	14.43	3023	5
Mathematics	61043	92.56	40730	66.72	9368	15.34	8337	13.65
CBT								
	Total Number Sat		A1-C6		D7-E8		F9	
	N	%	N	%	N	%	N	%
English Lang.	2107	88.04	970	46.03	575	27.28	344	16.32
Mathematics	2058	86	559	27.16	457	22.2	1022	49.65
Source: ICTD, WAEC								



# Performance: Mean score

Analysis of Multiple-Choice Question (MCQ) component of English Language and General Mathematics

Group Statistics for English Language and General Mathematics			
	MODE	MEAN	SD
ENG SCORES	PBT	55.89	10.948
	CBT	48.60	10.569
MATH SCORES	PBT	35.68	9.356
	CBT	22.70	9.029

Candidates who took English Language via PBT scored a mean of 55.89, compared to 48.60 for CBT, reflecting a notable difference of 7.30. In General Mathematics, the mean score was 35.68 for PBT and 22.70 for CBT, with a substantial gap of 12.98.

Independent Samples Test					
		t	sig	Sig (2 -tailed)	Mean difference
ENG SCORES	Equal variances assumed	28.428	0.49	<0.001	7.296
	Equal variances not assumed	29.377		<0.001	7.296
MATH SCORES	Equal variances assumed	60.710	0.79	0.00	12.983
	Equal variances not assumed	62.768		0.00	12.983
Independent Samples Effect Sizes					
				Point estimate	
ENG SCORE	Cohen's d			0.667	
	Hedges' correction			0.667	
	Glass's delta			0.690	
MATH SCORE	Cohen's d			1.389	
	Hedges' correction			1.389	
	Glass's delta			1.438	

Sig. (2-tailed) < .001 observed in English Language and Sig. (2-tailed) < .000 in General Mathematics indicates that statistically significant difference exists in the scores between both test modes.

Overall, candidates who sat for the PBT performed significantly better than those in the CBT mode.

# Incidence of Malpractice

## Research Question 2: How does the examination mode influence the rate of examination malpractice?

Findings indicate that the use of CBT was very effective in reducing examination malpractice compared to the traditional pencil-and-paper mode.

Only 19 malpractice cases (0.85%) were recorded in the CBT mode, compared to 2,498 cases (3.98%) in the PBT mode. Fewer types of examination malpractice occurred in the CBT mode (4) compared to the PBT mode (9).

S/N	Type of Malpractice	PBT: No. Involved (% of Total Sat)	CBT: No. Involved (% of Total Sat)
1	Bringing Foreign Materials into the Examination Hall	127 (0.20%)	5 (0.22%)
2	Caught with Mobile Phones	1,392 (2.22%)	12 (0.54%)
3	Irregular Activities Inside or Outside the Examination Hall	164 (0.26%)	– (0.00%)
4	Collusion	738 (1.17%)	2 (0.09%)
5	Impersonation	224 (0.36%)	1 (0.04%)
6	Foreknowledge	– (0.00%)	– (0.00%)
7	Leakage	– (0.00%)	– (0.00%)
8	Mass Cheating	– (0.00%)	– (0.00%)
9	Insult/Assault on Examination Officials	6 (0.01%)	– (0.00%)
10	Contravention of Instructions to Candidates	2 (0.003%)	– (0.00%)
11	Multiple Registration	14 (0.02%)	– (0.00%)
12	Destruction of Exhibits	– (0.00%)	– (0.00%)
13	Posting/Receiving Live Questions/Answers Online	1 (0.002%)	– (0.00%)



Feature	Paper-Based Test (PBT)	Computer-Based Test (CBT)
Difficulty	Wide range; many very easy, many difficult (some with $b=1.414$ ).	Wide range; many <i>extremely</i> easy items, also very difficult items.
Discrimination ('R')	Mixed: some good, many very poor (near zero).	Major Weakness: Generally, very low; several negative or near-zero 'R' values.
Discrimination ('a')	Highly variable: many extremely low (0.097), many extremely high (6.000), some moderate.	Major Weakness: Generally low to modest; few strong discriminators.
Guessing ('c')	Critical Weakness: Many items with extremely high 'c' (e.g., 0.700).	Better than PBT; mostly reasonable, a few on the higher side but not as extreme.
Key Concerns	Pervasive low 'a'/high 'c' combinations; 'a'=6.000 items; poor 'R' for some.	Pervasive low discrimination ('R' & 'a'); negative 'R' values; some extremely easy/hard items that don't discriminate.

# Psychometric Properties

## Research Question 3: What impact does the examination mode have on the psychometric properties of test items?

A post-test item analysis was carried out on the multiple-choice sections of English Language and General Mathematics. The comparison of the item parameters reports for English Language indicates that while both examination mode showed wide spread of difficulty parameter, PBT had a recurring 'b' value of 1.414 for many difficult items, often tied to extreme 'a' values. In some cases, extremely low 'a' values (0.097) and others problematically high 'a' values (6.000) on the test.

# Psychometric Properties cont'd

While many items had negative ‘b’ values in PBT, several items on CBT had predominantly positive 'b' values, confirming higher difficulty. There were a cluster of items a with 'b' value of 1.538, often linked with very low 'a' parameters, suggesting these 'b' values might be poorly estimated or constrained and represent difficult items.

*The empirical analysis of the psychometric properties of test items from both examination modes and across the selected subjects revealed significant differences.*

Summary of General Mathematics Item Parameters for All Calibrated Items

Feature	Paper-Based Test (PBT)	Computer-Based Test (CBT)
Difficulty	Generally easier; wider range, (some with b=1.538). .	Generally harder; skewed towards difficult items.
Discrimination	More variable: some very low 'a', some very high 'a', some good. Classical 'R' mixed.	More modest & consistent 'a'; generally lower classical 'R'. Contains items with negative/very low 'R'.
Guessing ('c')	Major weakness: Several items with extremely high 'c' values.	Strength: 'c' values generally stable and reasonable.
Key Concerns	High guessing, very low discrimination on some items, potentially unstable 'a' and 'b' estimates for those.	Overall higher difficulty, low classical discrimination for many, a few highly dysfunctional items (e.g., negative 'R').

# Validity & Reliability of the Test

**Research question 4: How does the examination mode affect the validity and reliability of the test?**

Test validity was ensured through rigorous standardization by WAEC's Test Development Division, aligned with the approved curriculum and global best practices.





# Validity & Reliability of the Test

Cronbach's alpha was used to assess the reliability of candidates' responses in both subjects.

Reliability analysis showed excellent internal consistency for PBT, with values of 0.967 (English) and 0.959 (Math), while CBT also demonstrated good reliability at 0.882 (English) and 0.879 (Math). The slightly lower CBT scores may be attributed to factors such as unfamiliarity with the format, interface differences, or test anxiety.

*Validity and reliability results showed no significant effect of test mode.*

*Reliability estimates of the assessment instruments on both test modes*

	Mode	Cronbach's Alpha	N of items
ENGLISH	PBT	0.967	80
	CBT	0.882	80
MATH	PBT	0.959	50
	CBT	0.879	50

# Final Results

Your the final enfigs compend piris tout furtiga an prosebdiaris of there notured therstes that deesig, subliction experimend and blets and lm the esscending uz firmo, the ploed of to scome thele techmehuok.



210%

This is lineals met carzed firted an secong of rextion for unatie in the (rleuas of experioned lnta fons in the ceprerend of shaed orcurtion and respades of the socuasiory scourt on decranent or carlor.



70%

This is is and the coollctedces ent euppertesly or fencolor gas of the etoppated ter) scuride ser and of the destines and intente for the dealer's lineacality nrfirevellesity.



60%

This is and) teccant an aconcleration foveransceccidant)scorties on this securidy speed foveres on and and dlcoves in the matorcal beas in. The repeas ts oth the odens ene dleg.

Experiment successfully demonstrates the effectiveness of the new method.

# Conclusion

## Key Findings: PBT vs CBT in WASSCE 2024 (PC - Second Series)

The Study compared performance, malpractice rates, and test reliability between PBT and CBT modes.

- **Performance Outcomes:**
  - **PBT candidates outperformed** CBT candidates significantly.
  - Large effect sizes confirmed the practical significance of the difference.
- **Malpractice Incidents:**
  - **Lower incidence in CBT**, highlighting enhanced test security.
- **Reliability Metrics:**
  - Acceptable reliability in both modes.
  - Slightly **higher internal consistency** observed in PBT.
- **Psychometric Validity:**
  - All candidates answered the same items.
  - Item-level analysis confirmed **validity, reliability, and fairness** across modes.
- **Overall**
  - **Trade-off exists:** Higher scores in PBT vs. better security in CBT.
  - Supports **CBT as a future-oriented solution** despite current performance gaps.

# Recommendations

## Conduct Further Research

Researchers and other stakeholders should conduct further research on the effect of test modes on performance of candidates.

## Adopt Phased CBT Approach

WAEC and other Examining bodies should adopt a phased or blended approach for CBT implementation.

## Invest in Infrastructure and Access

Government and other Major stakeholders should invest in Infrastructure and Access.

## Encourage Digital Literacy Training

Introduction and encouragement of Digital Literacy Training by Government and other stakeholders.

## Improve CBT Interface & Support

CBT Interface Improvement and Support Systems by Developers.

## Monitor & Evaluate Test Mode Impact

Government, Examining Bodies and Educational Institutions should Monitor and Evaluate Test Mode Impact Regularly.

## Continuous Public Sensitization

Continuous Public Sensitization and Stakeholder Engagement by Examining Bodies.





**THANK YOU  
FOR LISTENING**

