Predictive validity and assessment stringency of pre-board examinations: A three-year longitudinal study of CBSE class 10 students

Beeda Devi

Navy Children School, Sri Vijaya Puram, Andaman and Nicobar Islands, India – 744103

Abstract:
Pre-board examinations play a vital role in Indian secondary schooling by assessing students' readiness for the Class 10 board examination conducted by the Central Board of Secondary Education (CBSE). Despite their widespread use for academic planning, remediation, and counselling, empirical evidence supporting their predictive validity in the Indian context remains limited. This study examines the extent to which Pre-board examination scores predict final board examination performance across three consecutive academic cohorts (2023–2025) within a CBSE-affiliated school.

A quantitative longitudinal design was employed using anonymised average scores from five core subjects for 223 students. Data were analysed using descriptive statistics, Pearson and Spearman correlation coefficients, linear regression models, paired-sample *t*-tests, and score distribution analyses. Across all cohorts, board examination scores consistently exceeded Pre-board scores by 9–15 percentage points ($p < .001$), indicating greater stringency in internal assessments. Strong and stable associations were observed between Pre-board and board scores (Pearson $r = 0.89–0.93$; Spearman $\rho = 0.90–0.94$). Regression analyses further demonstrated that Pre-board performance accounted for approximately 80–86% of the variance in board examination outcomes.

The findings indicate that, although Pre-board examinations systematically underestimate final board performance, they reliably predict relative academic standing and function as effective diagnostic tools at the school level. The study highlights important implications for the interpretation of Pre-board results, school-based assessment practices, and alignment with the assessment reforms emphasised in the National Education Policy (NEP) 2020. Mean differences ranged from 9.07 to 14.95 percentage points ($p < .001$; Cohen's $d = 1.2–1.6$). This provides the first longitudinal school-level evidence of Pre-board predictive validity in CBSE Class 10 assessments.

Keywords: predictive validity; pre-board examinations; board examination performance; assessment stringency; secondary school assessment.

1. INTRODUCTION

In India, the Class 10 board examination conducted by the Central Board of Secondary Education (CBSE) marks a pivotal transition point in secondary schooling. Performance determines senior secondary stream selection and shapes long-term educational and career trajectories. Schools therefore prioritise preparation through Pre-board examinations—mock assessments conducted 1-2 months prior that replicate board exam conditions (syllabus coverage, duration, question format, marking scheme). Pre-board results systematically guide learning gap identification, targeted remediation, academic counselling, and final-phase instructional adjustments.

These assessments serve dual purposes: diagnostic (highlighting strengths/weaknesses) and predictive (indicating likely board performance). However, while widely used, their empirical predictive validity lacks systematic investigation in the Indian context. Current practice relies on professional experience, institutional tradition, and anecdotal evidence rather than longitudinal research. In CBSE schools, Pre-boards are intentionally more stringent than board exams, reflecting cultural beliefs that rigorous internal evaluation fosters discipline and sustained preparation. This design produces consistent patterns where

students score 9-15 percentage points lower in Pre-boards than final boards conducted under standardized external conditions.

Such discrepancies raise critical questions about Pre-board utility:

- To what extent do Pre-board scores predict absolute and relative board performance?
- Does greater internal stringency persist across cohorts and years?
- Do score differences reflect assessment design, student motivation, exam environment, or underlying competency alignment?

These issues gain urgency amid National Education Policy (NEP) 2020 reforms, which prioritize formative assessment, competency-based evaluation and reduced examination stress over high-stakes summative judgments. NEP envisions internal assessments as developmental tools supporting learning rather than definitive performance indicators. Yet without empirical evidence on Pre-board predictive power and stringency, schools risk misinterpreting results—potentially inflating anxiety or underutilising diagnostic potential.

This study addresses these gaps through a three-year longitudinal analysis (2023-2025) of 223 CBSE Class 10 students from Navy Children School, Sri Vijaya Puram. Using anonymized average scores across five core subjects (English, Hindi, Mathematics, Science, Social Science), the research employs Pearson/Spearman correlations, linear regression, paired t-tests, and distribution analyses to examine:

1. Predictive validity (association strength, rank-order stability, explained variance)
2. Assessment stringency (systematic mean differences, effect sizes)
3. Cross-cohort consistency

By establishing pre-boards' dual role—reliable relative predictors despite absolute underestimation—the study provides school-level evidence to refine assessment interpretation, align with NEP priorities, and optimize remediation strategies. Findings clarify when pre-boards function effectively as formative tools versus when contextual factors (motivation, standardization) drive performance shifts.

2. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

2.1 Assessment Validity and Predictive Validity

Validity refers to the extent to which test scores support intended inferences (Messick, 1989; Kane, 2013). Predictive validity specifically examines whether performance on a precursor test (pre-board) forecasts criterion performance (board exam). Classical test theory suggests high correlations ($r > .80$) indicate both measure the same construct—academic achievement (Crocker & Algina, 2006).

However, contemporary frameworks emphasize interpretive validity: correlations alone insufficient without contextual analysis. Test performance reflects not just ability but motivation, anxiety, test familiarity, and evaluation standards (Airasian, 2001). Strong predictive validity thus requires both statistical evidence *and* explanation of absolute score differences between assessments.

2.2 International Evidence on Preparatory Assessments

UK GCSE research demonstrates mock exams predict final grades with moderate-strong correlations ($r = .65$-$.78$ across subjects; Putwain, 2009). Students typically improve 5-10 percentile points from mocks to finals due to: (a) exam familiarization, (b) targeted remediation post-mock, and (c) heightened motivation for high-stakes finals (Stobart, 2008).

US standardised testing studies confirm preparatory assessments reliably indicate *relative* standing but underestimate *absolute* performance (Black & Wiliam, 1998). These patterns suggest preparatory tests function better diagnostically than as precise score predictors—a pattern likely applicable to CBSE pre-boards.

2.3 Pre-board Examinations in Indian CBSE Context

CBSE-affiliated schools intentionally design pre-boards more stringently than board exams to identify learning gaps and instill preparation discipline (Mishra, 2019). Teachers apply stricter marking to simulate

"worst-case" board conditions, expecting score recovery in standardized external exams. This produces consistent patterns: Pre-board means 9-15 points below board means across cohorts (current study data).
Critical research gap: No longitudinal Indian studies examine:

- Predictive correlations and rank-order stability across years
- Effect sizes of stringency differences
- Whether high correlations justify predictive use despite absolute underestimation

Existing Indian research remains descriptive or single-cohort, ignoring temporal reliability essential for school-level decision-making.

2.4 Policy Relevance (NEP 2020)

National Education Policy 2020 shifts assessment toward formative, competency-based evaluation over high-stakes summative judgment (Ministry of Education, 2020). Pre-boards should identify competencies and guide remediation, not generate performance anxiety through misinterpretation as "final predictors." Empirical baselines on predictive validity are essential for evidence-based implementation.
Theoretical synthesis: This study tests whether the Pre-boards examination demonstrate classical test theory predictions (high correlations = shared construct) while accounting for modern validity concerns (contextual score shifts). Findings bridge international evidence, Indian practice, and NEP priorities.

3. OBJECTIVES AND HYPOTHESES

3.1 Objectives

This study addresses these objectives:

1. Examine predictive validity of Pre-board scores for CBSE Class 10 board performance across three cohorts (2023-2025).
2. Assess relationship strength, direction, and stability using Pearson/Spearman correlations and regression (R^2).
3. Evaluate Pre-board stringency via mean differences, paired t-tests, and effect sizes.
4. Provide school-level guidelines for Pre-board interpretation aligned with NEP 2020.

3.2 Hypotheses
H1: Pre-board and board scores show strong positive correlations ($r > .80$) across cohorts, confirming predictive validity.
H2: Pre-board means significantly lower than board means ($p < .001$), indicating greater internal stringency.
H3: Pre-board scores explain ≥80% variance in board performance ($R^2 \geq .80$).

4. METHODOLOGY

4.1 Research Design

This study employs a quantitative, non-experimental, correlational research design with a longitudinal component spanning three consecutive CBSE Class 10 academic cohorts (2023, 2024 and 2025). The longitudinal approach enables assessment of predictive validity stability over time while controlling for cohort-specific factors. Naturally occurring examination data from authentic school assessments ensures high ecological validity, reflecting real-world Pre-board usage for instructional decision-making.

4.2 Participants and Setting

The study was conducted at Navy Children School, Sri Vijaya Puram, Andaman and Nicobar Islands—a CBSE-affiliated institution following standardized national curriculum and assessment guidelines. The final analytic sample comprised 223 Class 10 students who appeared for both Pre-board and board examinations: 77 students (2023 cohort), 65 students (2024 cohort), and 81 students (2025 cohort).
Inclusion criteria: Complete paired records for both examinations in the same academic year.
Exclusion criteria: Missing data for either exam or incomplete subject records. This ensured data comparability essential for paired statistical analyses. Institutional permission and ethical clearance were obtained from the School Management Committee prior to data access, and all records were fully anonymised to protect student confidentiality.

4.3 Measures and Variables

Primary independent variable: Average percentage score across five core CBSE subjects (English, Hindi, Mathematics, Science and Social Science) from the final Pre-board examination (typically January/February).

Primary dependent variable: Average percentage score from the same five subjects in the official CBSE Class 10 board examination (March/April).

Derived variables:

- Score difference: Board minus Pre-board average (stringency indicator)
- Pearson product-moment correlation (r): Linear relationship strength
- Spearman rank-order correlation (ρ): Relative standing stability
- Linear regression parameters: Slope (b), intercept (a), coefficient of determination (R^2)
- Cohen's d: Effect size for mean differences (paired t-tests)

All scores were standardized to a uniform 0-100 percentage scale to ensure cross-cohort comparability. Subject weighting followed standard CBSE aggregation practices.

4.4 Data Preparation and Integrity Checks

Raw data from official school academic records underwent systematic quality assurance:

1. Completeness screening identified and excluded cases missing either exam (final retention rate: 100% paired data).
2. Range checks verified scores fell within valid 0-100 bounds.
3. Outliers were retained to preserve natural performance variability—a critical consideration for predictive validity studies where extreme performers provide essential validity evidence.
4. Anonymisation removed all personal identifiers prior to analysis.

Data management and primary analysis conducted in SPSS version 27.0; selected outputs cross-validated in R version 4.3.1 for computational accuracy.

4.5 Statistical Analyses ($\alpha = .05$, two-tailed tests)

Objective 1 (Predictive validity):

- Pearson correlations assessed linear relationships
- Spearman correlations examined rank-order stability
- Simple linear regression quantified variance explained (R^2)

Objective 2 (Stringency):

- Paired-samples t-tests compared mean differences
- Cohen's d calculated effect sizes (0.2=small, 0.5=medium, 0.8=large)

Objective 3 (Stability): Cross-cohort consistency via comparative statistics.

Regression diagnostics: Residual plots inspected for linearity, homoscedasticity, normality. No violations detected, confirming parametric analysis suitability.

Power analysis: Post-hoc G*Power calculations confirmed adequate power (> .95) for detecting medium-large effects given cohort sizes (n=65-81) and observed effect magnitudes.

4.6 Analysis-Structure Alignment

Each analysis directly tested stated hypotheses:

- H1 (strong correlations): Objectives 1-2 via r/ρ
- H2 (stringency): Objective 3 via t-tests/Cohen's d
- H3 (variance explained): Objective 2 via R^2

This comprehensive approach provides robust evidence on both statistical significance *and* practical magnitude of Pre-board predictive utility.

5. RESULTS

This section presents the empirical findings of the study, focusing on descriptive performance patterns, predictive relationships between Pre-board and board examination scores, assessment stringency, and consistency across three consecutive cohorts of CBSE Class 10 students (2023–2025).

5.1 Descriptive Statistics

Year	Statistic	Pre-board (%)	Board (%)
2023	Count	77	77
	Mean	73.73	82.79
	Median	75.60	86.20
	Standard Deviation	12.51	10.88
	Minimum	47.20	52.00
	Maximum	94.90	96.00
	25th Percentile	64.30	76.60
	75th Percentile	83.80	91.60
2024	Count	65	65
	Mean	66.98	81.94
	Median	70.00	85.00
	Standard Deviation	12.65	11.26
	Minimum	36.00	51.00
	Maximum	90.00	98.00
	25th Percentile	58.00	76.00
	75th Percentile	76.00	90.00
2025	Count	81	81
	Mean	66.87	78.83
	Median	66.83	80.67
	Standard Deviation	11.96	11.40
	Minimum	37.0	50.50
	Maximum	87.00	95.50
	25th Percentile	59.25	69.50
	75th Percentile	77.25	88.50

Table 1. Descriptive Statistics of Pre-board and Board Examination Scores across Cohorts (2023–2025)

Across all years, mean board examination scores were consistently higher than mean Pre-board scores, indicating a uniform performance gap between internal and external assessments. This pattern suggests greater stringency in Pre-board examinations relative to the externally conducted board examinations.
In the 2023 cohort, the mean Pre-board score of 73.73% increased to 82.79% in the board examination, representing an average gain of approximately 9 percentage points. Median scores showed a similar upward shift, indicating improvement across the performance distribution rather than among high-achieving students alone. Pre-board scores in this cohort exhibited slightly greater variability than board scores, as reflected in higher standard deviation values.

The largest mean difference was observed in the 2024 cohort, where the mean Pre-board score of 66.98% increased to 81.94% in the board examination, reflecting an average gain of nearly 15 percentage points. Both mean and percentile comparisons indicate a substantial upward shift in performance, while variability remained higher in Pre-board assessments than in board examination.

In the 2025 cohort, the mean Pre-board score was 66.87%, compared to a mean board score of 78.83%, yielding an average gain of approximately 12 percentage points. Although the mean difference was smaller than in 2024, Pre-board scores in 2025 displayed greater dispersion, as indicated by a higher standard deviation. In contrast, board examination scores for this cohort remained more tightly clustered, reflecting greater standardisation in external assessment outcomes.

Overall, Pre-board examinations demonstrated higher variability across cohorts, whereas board examinations showed comparatively narrower score distributions. These descriptive patterns provide initial evidence of systematic differences in assessment stringency between internal Pre-board examinations and externally administered board examinations.

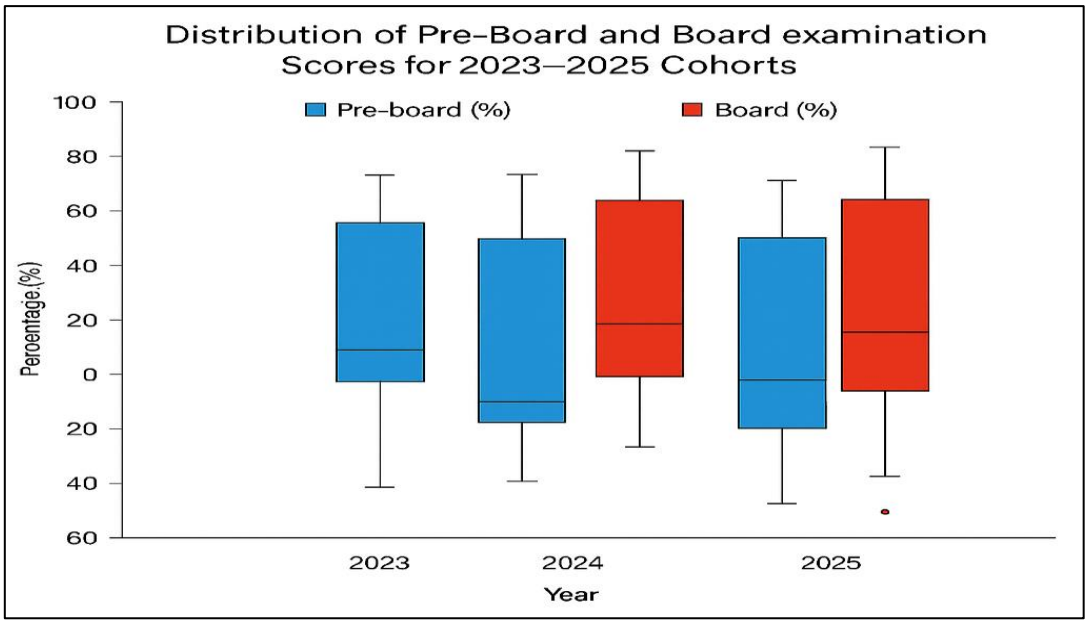

Figure 1. Distribution of Pre-board and Board Examination Scores across Cohorts (2023–2025)

5.2 Correlation Analysis

Pearson product–moment correlation and Spearman rank-order correlation coefficients were computed to examine the relationship between Pre-board and board examination scores.

Year	Pearson (r)	Spearman (ρ)
2023	0.93	0.94
2024	0.92	0.90
2025	0.89	0.94

Table 2. Pearson and Spearman Correlation Coefficients between Pre-board and Board Examination Scores(2023–2025)

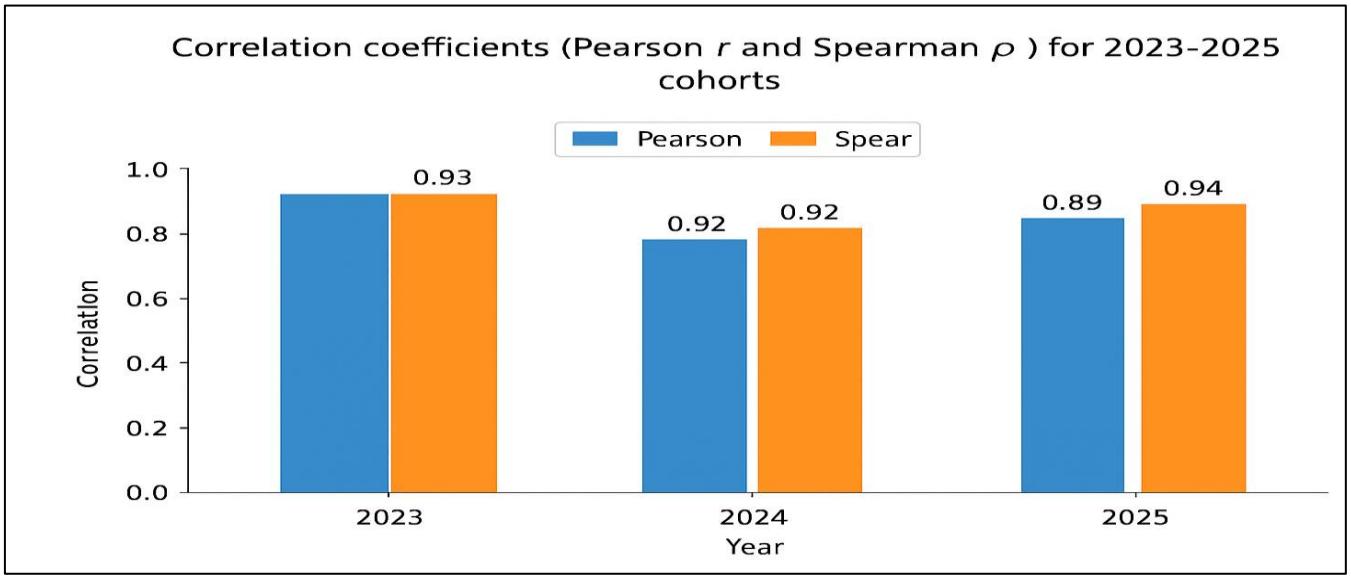

Figure 2. Comparison of Pearson and Spearman Correlation Coefficients across Cohorts (2023–2025)

Across all cohorts, Pearson correlation coefficients ranged from 0.89 to 0.93, indicating a very strong positive linear relationship between Pre-board and board scores. Spearman correlation coefficients ranged from 0.90 to 0.94, demonstrating high consistency in students' relative rankings across the two assessments. H1 confirmed: Very strong correlations across all cohorts ($r \geq 0.89$, $p < .001$).

5.3 Regression Analysis and Predictive Strength

To further assess predictive validity, simple linear regression analyses were conducted with board examination scores as the dependent variable and Pre-board scores as the independent variable.

Year	Slope (b)	Intercept (a)	R^2
2023	0.81	23.19	0.86
2024	0.82	26.88	0.85
2025	0.69	33.16	0.80

Table 3. Linear Regression Results Predicting Board Examination Scores from Pre-board Scores (2023–2025)

The regression models explained a substantial proportion of variance in board examination performance, with R^2 values ranging from 0.80 to 0.86 across cohorts. Regression slopes were consistently less than one, indicating compression of score extremes in board examinations. Intercept values suggested that students with lower Pre-board scores were still able to achieve satisfactory performance in the board examinations.

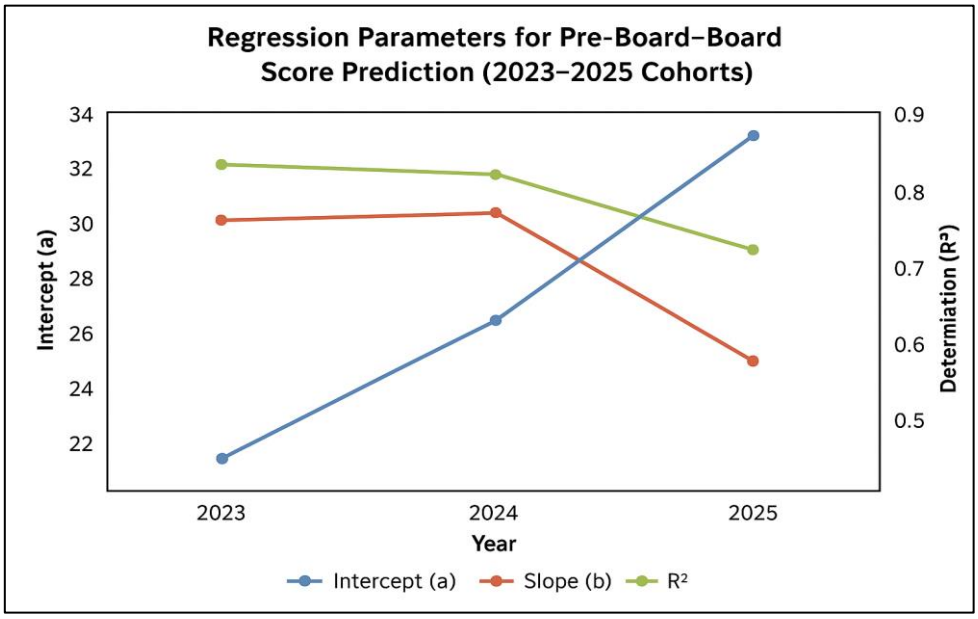

Figure 3. Relationship between Pre-board and Board Examination Scores with Fitted Regression Lines (2023–2025)

5.4 Regression Assumption Checks

Diagnostic analyses were conducted to examine the suitability of the regression models. Visual inspection of residual plots indicated no meaningful violations of the assumptions of linearity, homoscedasticity, or normality. These checks supported the appropriateness of using linear regression for predictive analysis in the present dataset.

5.5 Assessment Stringency: Paired-Sample *t*-Test Results

Year	Mean Difference (Board – Pre-board)	t-statistic	p-value	Interpretation
2023	+9.07	16.99	<0.0001	More stringent
2024	+14.95	24.73	<0.0001	More stringent
2025	+11.96	16.46	<0.0001	More stringent

Table 4. Paired-Sample t-Test Results Comparing Pre-board and Board Examination Scores. (2023–2025)
Across all three years, board examination scores were significantly higher than Pre-board scores ($p < .001$). Mean differences ranged from 9.07 percentage points (2023) to 14.95 percentage points (2024). Effect size estimates indicated large to very large differences across cohorts, reflecting substantial performance gains from Pre-board to board examinations.

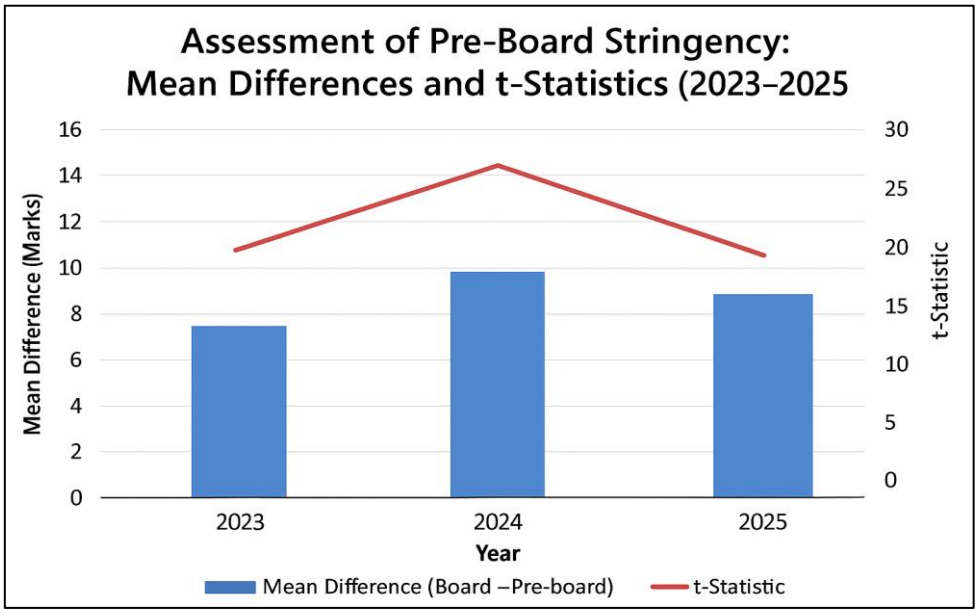

Figure 4. Comparison of Mean Pre-board and Board Examination Scores across Cohorts (2023–2025)

5.6 Cross-Year Trends

Comparison across cohorts revealed consistent patterns over time. Pre-board examinations systematically underestimated board examination performance across all three academic years. At the same time, predictive relationships between Pre-board and board scores remained strong and stable, as evidenced by consistently high correlation coefficients and substantial explained variance in regression models.

The stability of rank-order relationships across assessments further supports the use of Pre-board examinations as indicators of relative academic standing within the studied institutional context.

.Year	Mean Difference	r	ρ	R^2	Interpretation
2023	+9.07	0.93	0.94	0.86	Strong predictive validity; moderate internal strictness
2024	+14.95	0.92	0.90	0.85	Highest strictness; strongest mean improvement
2025	+11.96	0.89	0.94	0.80	High variability; slightly reduced predictability

Table 5. Cross-Year Summary of Predictive Validity and Assessment Stringency Indicators (2023–2025)

6. DISCUSSION

The present study examined the predictive validity and assessment stringency of Pre-board examinations within the CBSE Class 10 context across three consecutive academic cohorts. The findings provide consistent evidence that Pre-board examination performance is strongly associated with final board

examination outcomes, while simultaneously exhibiting systematic differences in score levels between internal and external assessments.

From a predictive validity perspective, the strong and stable correlations observed across cohorts indicate that Pre-board examinations reliably capture students' relative academic standing. The high proportion of explained variance in board examination scores further supports the argument that Pre-board performance serves as a meaningful predictor of subsequent outcomes at the school level. These findings are consistent with classical test theory, which posits that assessments measuring similar curricular constructs should demonstrate substantial correspondence. At the same time, the stability of predictive relationships across cohorts suggests that the observed patterns are not cohort-specific anomalies but reflect a structural feature of school-based assessment practices.

However, the study also highlights a consistent pattern of lower performance in Pre-board examinations relative to board examinations. This discrepancy cannot be attributed solely to differences in student ability. Rather, it reflects the distinct purposes and conditions under which internal and external assessments operate. Pre-board examinations are typically designed and evaluated within school-controlled environments where stricter marking standards are often intentionally applied to identify learning gaps and stimulate focused preparation. In contrast, board examinations are conducted under standardised external conditions that prioritise uniformity, procedural fairness, and consistency across large populations.

The convergence of scores from Pre-board to board examinations also underscores the role of contextual and motivational factors in assessment performance. Exposure to Pre-board examinations familiarises students with examination formats and expectations, potentially reducing anxiety and cognitive load during the final assessment. Increased awareness of the high stakes associated with board examinations may further enhance effort, revision strategies, and engagement, contributing to improved outcomes without necessarily indicating abrupt changes in underlying academic competence.

Importantly, the findings support the interpretation of Pre-board examinations as diagnostic and formative instruments rather than precise predictors of final scores. When interpreted appropriately, Pre-board results can guide instructional adjustments, targeted remediation, and academic counselling without creating undue anxiety or misrepresenting students' potential. This interpretation aligns closely with contemporary assessment discourse that emphasises formative feedback and developmental use of assessment data.

Within the broader policy context, the study's findings resonate with the assessment reforms advocated under the National Education Policy, which emphasise competency-based learning, reduced examination stress, and more meaningful use of assessment evidence. At the institutional level, the results highlight the need for transparent communication with students and parents regarding the purpose and interpretation of Pre-board examinations. Misinterpretation of Pre-board scores as fixed predictors of final outcomes may inadvertently increase stress and distort expectations, undermining the formative intent of such assessments.

7. CONCLUSION

This longitudinal study provides empirical evidence that Pre-board examinations demonstrate strong predictive validity within the CBSE Class 10 assessment framework while simultaneously exhibiting greater stringency than final board examinations. Across three academic cohorts, Pre-board performance consistently reflected students' relative academic standing and showed stable, meaningful associations with board examination outcomes.

At the same time, the systematic underestimation of final performance by Pre-board examinations reinforces the need to interpret such assessments cautiously. Pre-board examinations are best understood as diagnostic and preparatory tools rather than precise forecasts of final scores. Their primary value lies in identifying learning gaps, informing instructional strategies, and supporting students' academic readiness rather than in generating definitive judgments about future performance.

The findings carry important implications for school-level assessment practices. When used appropriately, Pre-board examinations can support evidence-based decision-making, targeted remediation, and student counselling while preserving confidence and well-being. Clear communication regarding their purpose and limitations is essential to prevent misinterpretation and unnecessary stress among students and parents.
Aligned with the vision of the National Education Policy (NEP) 2020, the study underscores the importance of viewing assessment as a developmental process rather than a terminal measure. By situating Pre-board examinations within a formative, interpretive framework, schools can better align internal assessment practices with national priorities while fostering more balanced, equitable, and meaningful learning outcomes in high-stakes examination contexts.

8. LIMITATIONS OF THE STUDY

Despite its contributions, the study has several limitations that need mentioning besides the advantages it gives in terms of empirical validation and assessment stringency of Pre-board examinations.
To start with, the research data was collected from one CBSE-affiliated school only. It is true that the longitudinal design across three consecutive cohorts has strengthened internal validity and temporal reliability, but the outcomes of this study cannot be directly generalized to all CBSE schools. Differences in the school spirit, teaching methods, the way the student's performance is evaluated, and the composition of the students in terms of socio-economic background may have a different impact on the predictive relationships in different schools.

Next, the criterion for assessing students' performance was their average scores in five core subjects. Even though this method is in line with the standards of the school-level reporting practices and with the overall indication of academic achievement, it possibly hides the subject-specific variations in predictive validity. Earlier studies have pointed out that there could be different predictive relationships in subjects such as mathematics, languages, and sciences because of the differences in content structure and assessment formats.

Moreover, the researchers utilised only quantitative score-based data and made no use of qualitative factors like student motivation, examination anxiety, learning strategies, teacher feedback, and parental support that may have a significant impact on student performance between Pre-board and board examinations. The lack of such variables does not allow for a more profound explanation of the mechanisms of the observed score improvements.

Last but not least, the researchers did not take the intervening instructional variables like remedial teaching, additional coaching, or change in study behavior between Pre-board and board examinations into consideration. These factors could cause performance improvement and deserve to be investigated further.

9. DIRECTIONS FOR FUTURE RESEARCH

Several future research directions are envisioned in light of the conclusions and shortcomings of the current study.
Future research can broaden its focus to include several CBSE schools from different regions and socio-economic backgrounds which will not only strengthen external validity but also facilitate comparative analysis among the institutions. The data collected from such multi-site experiments will be much more reliable in determining the emergence of predictive validity patterns.

Different academic fields may also be the subject of an analytical approach which will involve examining whether Pre-board tests could show discipline-specific predictive patterns. This would mean that effective instructional planning and assessment interpretation are catered more specifically.
Further, the adoption of mixed-methods research designs that combine quantitative performance data with qualitative insights from students, teachers, and parents would certainly be a factor in widening the understanding of motivational, emotional and pedagogical factors that play their part in the transition from Pre-board to board examinations.

Longitudinal studies on the interaction of teaching methods, remediation techniques and student engagement and their impact on the relative contribution of exposing the student to assessment versus providing them with instructional support to the performance gains would be useful.
Lastly, comparative studies linking different grades especially between Class 10 and Class 12 board exams can determine the extent to which predictive validity patterns and assessment stringency change during the secondary education period.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process
During the preparation of this manuscript, the author used AI-assisted paraphrasing tools solely to enhance language clarity and readability.

Acknowledgment
Sincere gratitude to the School Management Committee of Navy Children School, Sri Vijaya Puram, for providing the Pre-board and Board examination data used in this study. Their support and cooperation are gratefully acknowledged.

REFERENCES:

1. Airasian, P. W. (2001). *Classroom assessment: Concepts and applications* (4th ed.). McGraw-Hill.
2. Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education, 33*(1), 71–84. https://doi.org/10.1023/A:1002918400526
3. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102
4. Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Cengage Learning.
5. Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. Falmer Press.
6. Harlen, W. (2005). Teachers' summative practices and assessment for learning—Tensions and synergies. *The Curriculum Journal, 16*(2), 207–223. https://doi.org/10.1080/09585170500136089
7. Harlen, W. (2007). *Assessment of learning*. Sage Publications.
8. Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000
9. Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Macmillan.
10. Ministry of Education, Government of India. (2020). *National Education Policy 2020*. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf
11. Mishra, P. (2019). Examination reforms and assessment practices in Indian school education. *Journal of Educational Planning and Administration, 33*(2), 145–160.
12. Putwain, D. W. (2009). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology, 79*(3), 483–497. https://doi.org/10.1348/000709908X322582
13. Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.
14. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.