

Study of Existing Assessment Structure of CBSE Schools



Report by Oxford University Press, India.
© Oxford University Press 2022

Cover image: © arrowsmith2/Shutterstock
Inside image: © Chinnapong/Shutterstock

C O N T E N T S

Executive Summary 6

Introduction 8

Context 9

Aims and Objectives 10

Study Design 11

Research Methodology 11

Representative Data Sample 12

Underlying Key Assumptions 12

Key Evaluation Findings 13

Theme 1 13

Broad spectrum observation 13

Subject-specific observations 14

Theme 2 18

Broad spectrum observation 18

Subject-specific observations 19

Theme 3 21

Broad spectrum observation 21

Subject-specific observations 21

Theme 4 23

Broad spectrum observation 23

Subject-specific observations 25

Recommendations 28

Recommendation 1 29

Developing a comprehensive assessment framework

Recommendation 2 30

Framing a measurable blueprint for test construction

Recommendation 3 31

Ensuring rigour and quality in assessments

Recommendation 4 31

Specifying evaluation measures

Recommendation 5 32

Building assessment capacity

Conclusion 33

Appendix 35

Acknowledgements 36



Message from OUP



Fathima Dada,
Managing Director,
Education, OUP

On behalf of OUP (India), I would like to express how honoured and privileged we feel to have partnered with the Central Board of Secondary Education (CBSE) in a prestigious research project to explore the impact of assessment patterns followed by affiliated schools, with a special focus on grades VI to IX.

This study stems from the National Education Policy (NEP) 2020 thrust on attaining a desirable level of literacy and numeracy by ramping up the overall level-appropriate skills and knowledge of learners and building their proficiencies to match global (Programme for International Student Assessment [PISA]) standards, primarily through the application of appropriate pedagogy, curriculum and instruction design, and, finally, effective assessment design. This project has been conducted through a four-phase process, including Data Acquisition, Exploration, Data Analyses and Report Publication, where test papers for gateway subjects of mid- and end-term exams and report cards from student samples have been processed by independent consultants from the fields of assessments, psychometrics, and data analytics, and by experts from within OUP. The report incorporates assessment-related solutions and recommendations from experts and consultants to be able to understand the present trends and patterns in assessments and make informed, data-supported conclusions and suggestions.

OUP is very pleased to assist CBSE in conducting this research and envisages that the outcome will benefit institutions in mapping curriculum design and execution to the global requirement of outcome-driven learning, and thereby empower both teachers and learners in due course.

Message from CBSE



Manoj Ahuja,
Chairman, CBSE

I take this opportunity to appreciate the initiative taken by Oxford University Press India (OUP) to partner with Central Board of Secondary Education (CBSE) to conduct a much-needed research on assessment patterns being followed by schools affiliated with CBSE, with focus on grades VI to IX.

The CBSE, with more than 25,000 schools in its network, has always been committed to exploring ways of advancing the school teaching-learning system. Keeping in mind the salient requirements of NEP 2020, the purpose of the project is to ensure that learners are able to achieve expected literacy and numeracy levels, meet the stipulated core competence levels, and compete at national and international levels.

Since the most empirically effective way to achieve this is by directing all schools to engage in regular formative assessments to evaluate factual, conceptual, critical, and creative learning, it now becomes imperative to put in place a systematic study of existing performance-tracking methods that will help analyse to what extent they meet the desired learning outcomes and standards of learners.

This report, with its four-stage implementation—Data Acquisition, Exploration, Data Analyses, and Report Publication—will make for a clear understanding of how learners may benefit from the use of standardised, criteria-based assessment through term, half-yearly, and end-term assessments.

I would like to thank both OUP and CBSE for completing this first-of-its kind project in India, with a high degree of rigour and meticulousness as also an immense degree of commitment and dexterity. I do believe the end-users are to benefit largely from this endeavour.

EXECUTIVE SUMMARY

The Indian education system is currently reorienting itself from promoting “teaching to the test” to viewing tests/exams/assessments as a means of deeper learning. The new system of education being ushered in by the National Education Policy (NEP) 2020 aims at “holistic development” where each individual’s creativity and uniqueness is appreciated and given room to grow.

The fundamental changes in the approach towards assessments warranted by NEP 2020 brought into focus the need for understanding the existing assessment patterns being followed by schools affiliated with the Central Board of Secondary Education (CBSE), particularly for middle school. This shift in the education system’s priority, combined with the changes brought in by the CBSE’s realignment with the system of assessment and examination from its earlier Continuous and Comprehensive Evaluation (CCE) pattern, makes this research study crucial for incorporating any requisite changes as part of the larger shift in the overall system.

This study allows us access to the on-ground evidence of what patterns are being followed in the term-end exams being administered and to what extent they meet the objectives set out by the Board. It also provides deep insight into the attainment of the desired learning outcomes and level-appropriate skills and knowledge.

The study analysed the assessment patterns being followed by CBSE schools for grades VI to IX. It also examined the influence that exams have on achievement of the desired learning levels. The question papers of English, Hindi, Mathematics, and Science administered for the term-end exams of grades VI to IX for the Academic Year 2018-19 were studied along with the scoring patterns of the students.

The following figure shows the research methodology adopted for the study.

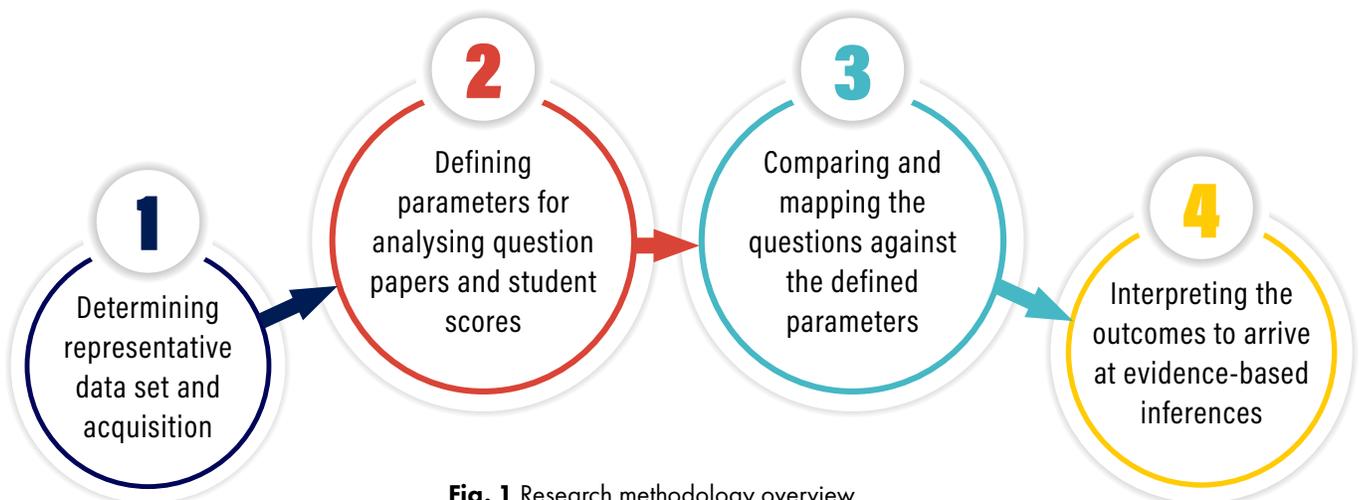


Fig. 1 Research methodology overview

The following are the **key findings** that emerged from the study:

- Partial uniformity was observed in the design of question papers across subjects and grades while implementing the core objectives of the Board regarding the assessment pattern.
- The attainment of literacy and numeracy skills across grades was also noted to be partial.
- The question papers were observed to be limited from a global standards perspective, resulting in the inability to check the proficiency levels of the students.
- The reporting of achievement of level-appropriate skills and knowledge by the students across grades was observed to be limited/incomplete due to unavailability of level-appropriate question papers in term-end exams.

The following **recommendations** are made basis the findings:

- ***Developing a comprehensible assessment framework*** that is competency-based and holistic and defines the assessment objectives, expected standards, and learning progressions for each grade and subject.
- ***Framing a measurable blueprint*** that details the test specifications such as number of items, coverage of topics, difficulty levels, percentage of questions pertaining to Higher Order Thinking Skills (HOTS).
- ***Ensuring rigour and quality in assessments*** through following of assessment principles of validity, reliability, fairness, usability, inclusivity, and relevance.
- ***Specifying evaluation measures*** that provide subject-specific broader goals at each grade, performance expectations and rubrics indicating the achievement criteria, and remedial measures to bridge the achievement gaps.
- ***Building assessment capacity*** for orienting teachers towards the use of assessment framework and blueprint, item development, reporting, and data management and interpretation.

Introduction

The education landscape in India is going through a paradigm shift due to the advent of the NEP 2020. This policy proposes the revision and revamping of all aspects of the education structure, including the assessment tools for evaluating the performances of the learners. With its focus on regular, formative, and competency-based assessments, and holistic measurement of learning, it has become vital to capture learning data and monitor the learning levels of the learners.

To bring about the NEP warranted metamorphosis in the assessment system, and to establish a robust mechanism to enable learners attain the desired learning outcomes, CBSE initiated the study of assessment patterns followed by the schools affiliated to the Board. This is particularly critical, as the grades below X are not formally structured.

This study aims to provide relevant empirical evidence to enable informed decisions at both the macro (policy) and operational levels to harmonise the existing assessment structure with NEP's overall assessment strategy.

The Board partnered with Oxford University Press (OUP) to conduct this study.

Context

The CBSE transitioned from its CCE pattern to a system of assessment and examinations from the Academic Year 2017-18 onwards, that saw the restoration of Board Examination/School-based Examination for Grade X. This resulted not only in revised assessment structure for Grade X but also in changes in the evaluation and assessment structure of middle school grades.

As per the new regime, each academic session is divided into two terms of 100 marks each. Figure 2 gives an overview of this system of assessment.

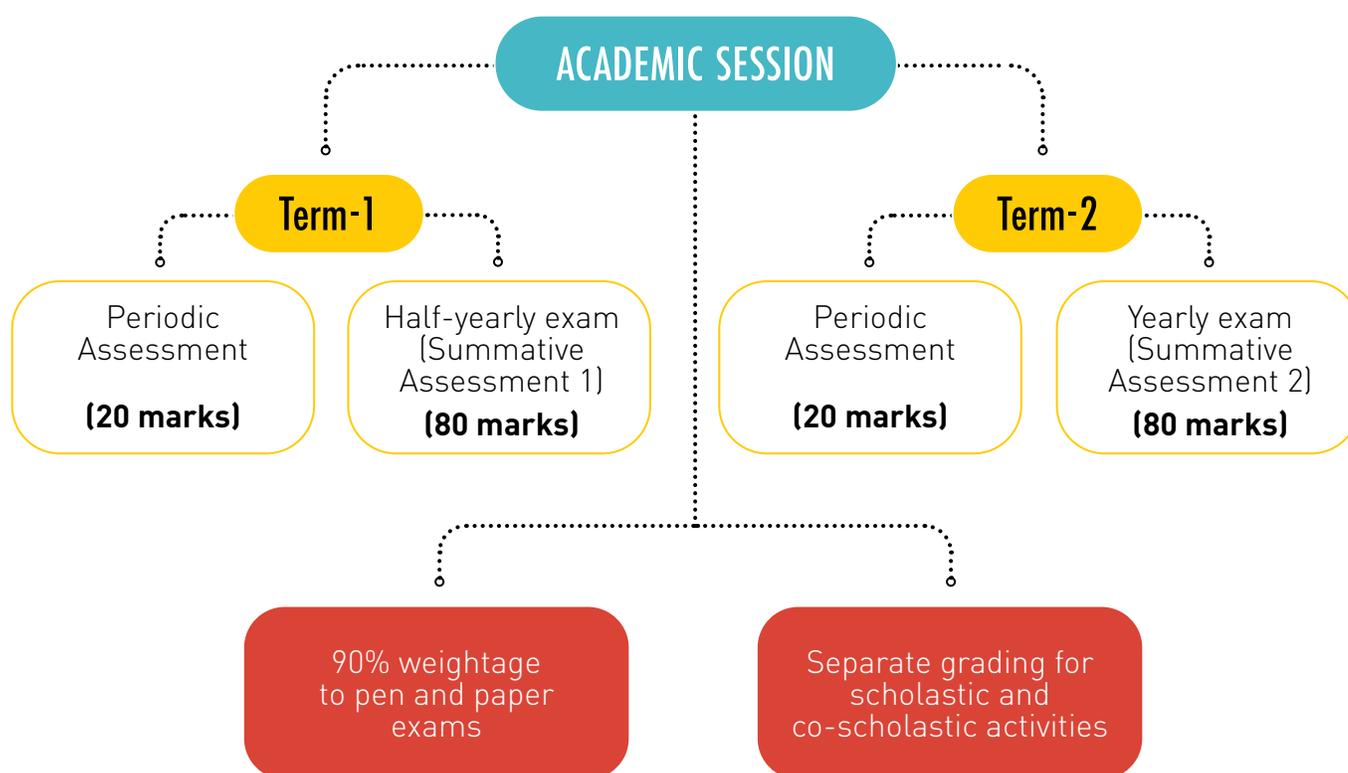


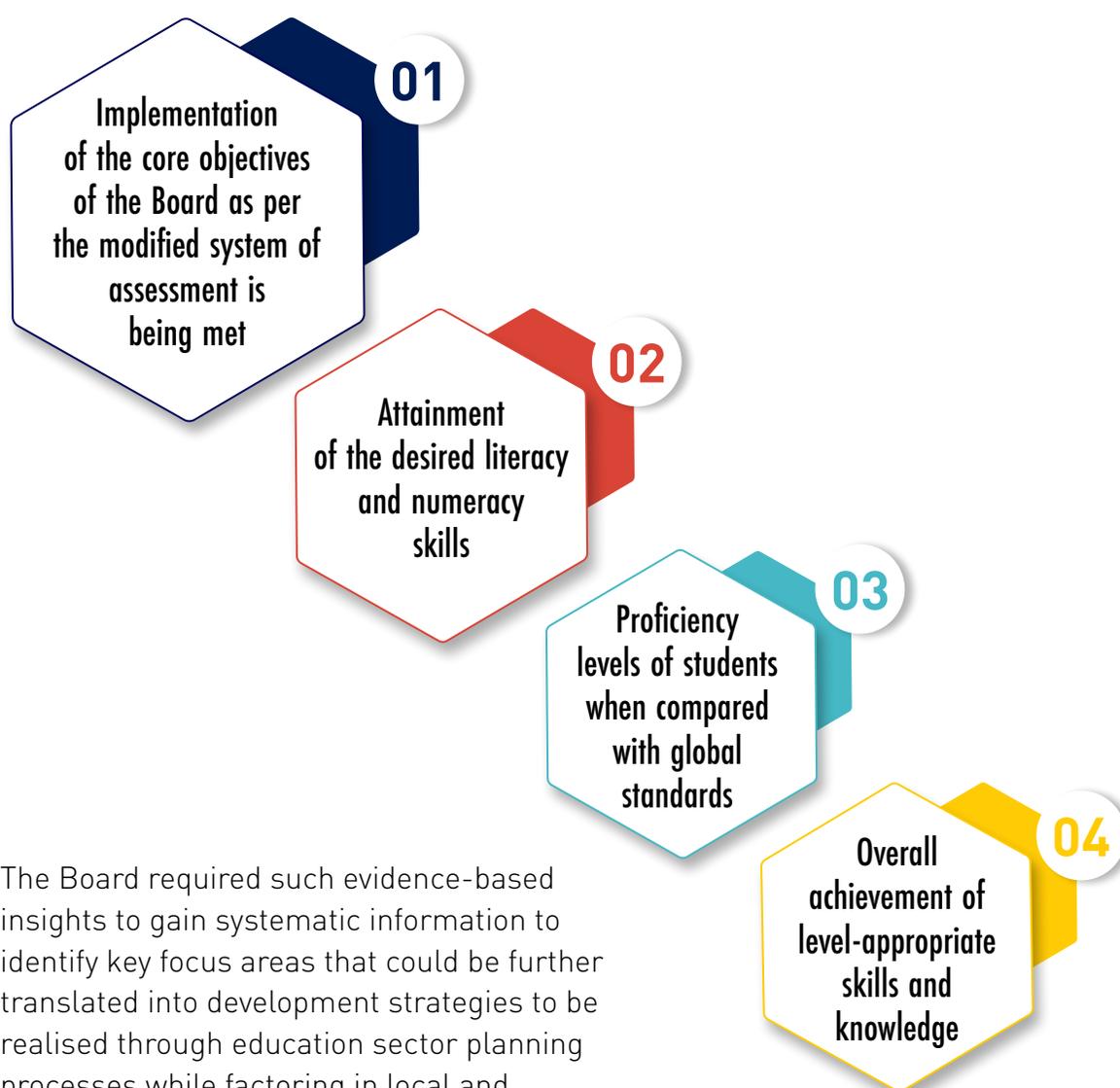
Fig. 2 Overview of CBSE’s modified system of assessment, examination and report card

The revision enabled the learners to progressively rise to the challenge of Grade X Board Examination from Grade VI. This is facilitated through the modified system of assessment and examination across all schools from Grade VI onwards.

Aims and Objectives

To measure the effectiveness of this transition, the Board wanted a mechanism that could help them assess the performance of middle school, across the gamut of its schools, as per the new system of assessment. The study conducted by OUP is aimed at examining the assessment patterns followed by the schools, affiliated with the Board, for grades VI to IX.

For a thorough understanding of the impact that exams have on achievement of the desired learning levels, the following themes were considered:



The Board required such evidence-based insights to gain systematic information to identify key focus areas that could be further translated into development strategies to be realised through education sector planning processes while factoring in local and regional needs.

Fig. 3 Themes of the study

Study Design

The study primarily involved analysis of acquired data by assessment specialists and working with domain experts in the field of psychometrics and data analytics. A two-pronged approach was exercised to analyse the assessment patterns followed by the schools. The assessment patterns were studied through understanding the question papers and scoring patterns of half-yearly and yearly exams conducted by the schools and their impact on attaining the desired learning levels.

Research Methodology

The study involved the following steps:

DATA ACQUISITION

- A heterogenous representative data set was defined
- Question papers and scores for English, Hindi, Mathematics, and Science of term-end exams (half-yearly and yearly) of grades VI-IX were acquired via the Board

EXPLORATION

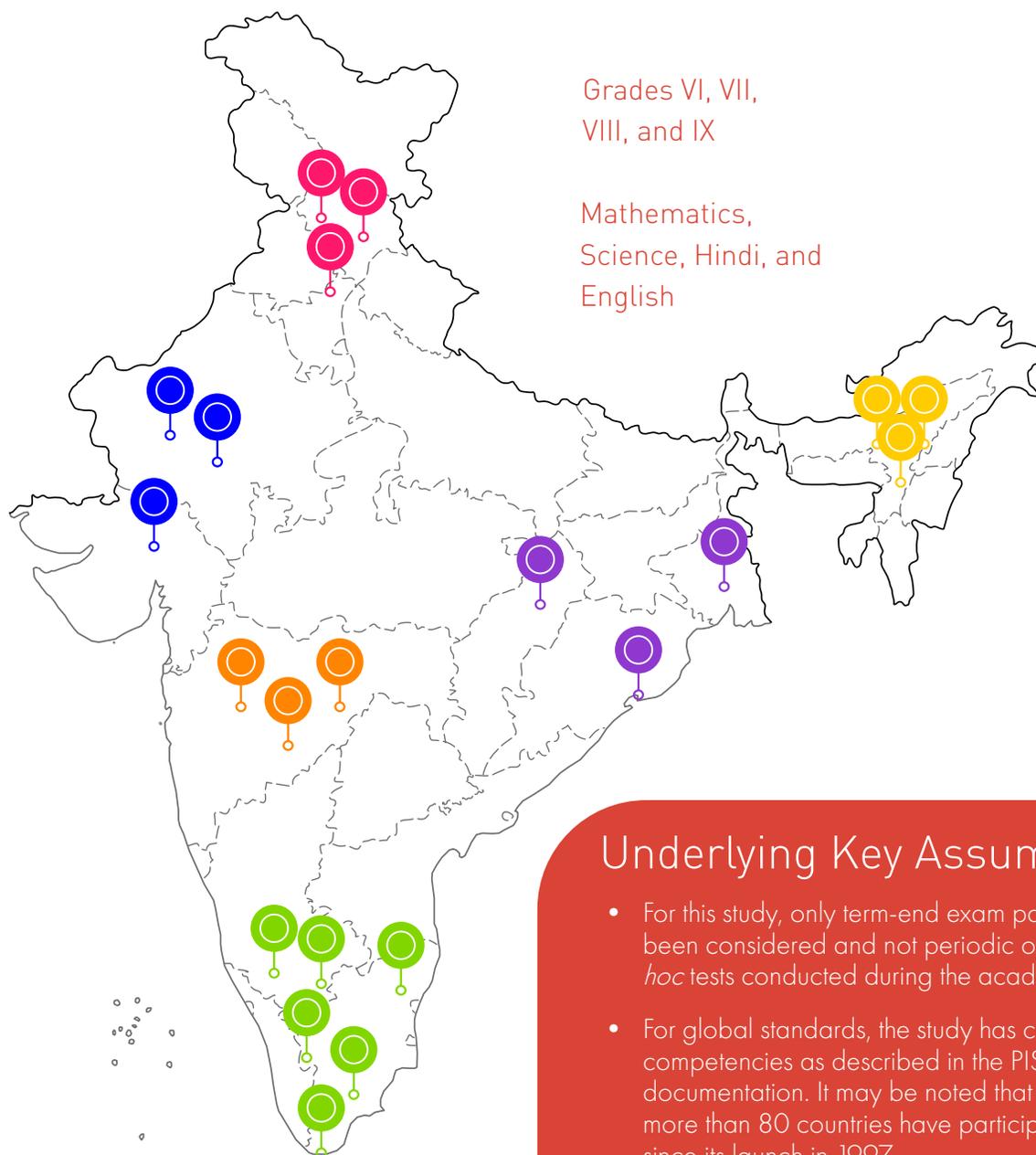
- Parameters for analysing question papers were defined, which included:
 - identifying what the exams were measuring
 - observing the coverage of different topics in the question papers
 - comparing test items/questions against the expected learning outcomes and against global assessment benchmarks such as Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS)
- Inputs received were organised, studied, compared, and mapped to the defined parameters of measurement; detailed interpretations specific to the subjects were recorded in templates

DATA ANALYSIS AND INSIGHTS

- Systematic and detailed statistical analysis was conducted
- Data modelling techniques were applied to the subject-specific templates and student scores were analysed and interpreted to arrive at informed, data-supported conclusions and insights

Representative Data Sample

For the study, representative data samples from four regions and eight Centres of Excellence (COEs) were considered.



Underlying Key Assumptions

- For this study, only term-end exam papers have been considered and not periodic or any other *ad hoc* tests conducted during the academic session.
- For global standards, the study has considered competencies as described in the PISA documentation. It may be noted that students from more than 80 countries have participated in PISA since its launch in 1997.
- Definition of literacy and numeracy skills considered in the study pertains to overall scores and the extent to which question papers adhered to desired learning outcomes for each level.
- Due to the lack of item-wise scores of students, the key findings of the study are based on the overall attainment of the scores and the quality of question papers administered in the exams.

Key Evaluation Findings

The analysis of question papers and corresponding scores of the students of term-end exams of Academic Year 2018-19 across schools clearly shows the areas that have robust examples of assessment tasks and the ones that need more attention for holistic evaluation. The research also points to the specific characteristics of each grade and subject assessment that aligns to the desired learning outcomes, attainment of literacy and numeracy skills, and the global standards.

THEME 1

Implementation of the core objectives of the Board as per the modified system of assessment—extent to which these are being met by schools

Theme 1 focuses on insights from a series of analyses of the data considered for this study. This theme focuses on factors related to the quality of question papers that are administered to the students in both the term-end exams. The broad goal of these analyses is to understand how far the exams conducted by the schools adhere to the core objectives of the Board for assessing English, Hindi, Mathematics, and Science across grades VI, VII, VIII, and IX.

Broad spectrum observation

There is partial uniformity observed in the design of question papers across subjects and grades.

The analysis has shown that on an average, 3-5% questions are above-grade level.

Extent of uniformity

- It is observed that the number of questions across the two term-end exams are almost the same.
- The questions across all subjects are mostly grade-appropriate in terms of mapping with the set learning outcomes.

Extent of non-uniformity

- The question papers do not contain questions that cater to all the relevant topics and learning outcomes for any specific level/grade with certain topics and learning outcomes not being covered at all.
- Largely, the questions are not designed in a manner that test the appropriate proficiency level for a grade even though they might map with the learning outcomes set for that grade.
- The question papers in general are skewed towards questions of Lower Order Thinking Skills (LOTS).

Subject-specific observations

ENGLISH

Extent of match

- Some of the question papers are of reasonably good standard and are grade-appropriate. They are well-designed, wherein most of the learning outcomes are measured and the difficulty level is appropriate, especially for grades VIII and IX.
- There is almost a uniform spread of questions at an overall level between the two term-end exams.
- The comprehension and grammar content for Grade VIII is relatively appropriate compared to other grades.
- There are some minimum number of questions falling under the Create level of Bloom's taxonomy.
- In comparison to other grades, grades VI and IX question papers comprise around 20% questions from HOTS.

Extent of mismatch

- Although the spread of questions at an overall level is uniform, the number of questions administered is diverse and ranges from a minimum number of 14 questions to a maximum number of 105 questions across grades.
- In general, the difficulty level of the question papers is 'easy' with limited questions pertaining to understanding of genre and text style analysis.
- Some of the question papers are very lengthy with too many questions carrying 1 or 2 marks.
- The comprehension and grammar content for grades other than Grade VIII leaves a lot of scope for improvement.
- For all the grades, a disproportionate percentage of about 90% of the questions are focused on prose, reading comprehension, and integrated grammar, leading to relative neglect of other topics.
- Most of the learning outcomes remain untested and primarily the lowest level of competencies is assessed for most of them.
- Questions pertaining to critical thinking, value-based analysis, decision-making, research and inquiry are minimal, accounting for only around 5% across all the grades.

The term-end exam papers' questions are largely focused on reading and writing skills with a variation in the spread of number of questions between reading and writing skills across grades (refer to Fig. 4).

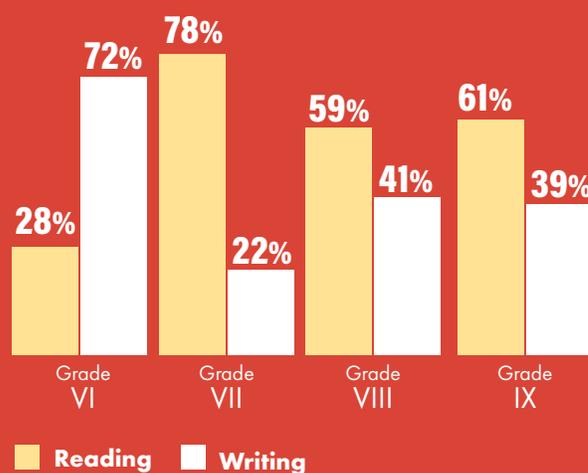


Fig. 4 Grade-wise question distribution between reading and writing skills

While **Grade VIII** has almost equal distribution between the two skills, in **Grade VI**, reading questions only constitute **28%** as against **72%** of writing questions. **Grade VII** has **78%** of the questions focusing on reading and **Grade IX**, again, has more focus on reading with **61%** questions.

Extent of match

- Several question papers are well-balanced in terms of having the levels of Bloom's taxonomy and assessing the grade-appropriate learning outcomes.
- It is noted that the spread of questions across both the term-end exams is equal in grades VI and VIII.
- About 90% of question papers are grade-appropriate.
- Spread of difficulty level: a good percentage of questions belong to "easy" category followed by "medium" category and about 10% from "difficult" category.
- Question papers across grades and schools include questions from creative thinking and about 14% questions from HOTS.

Extent of mismatch

- In several cases, the format and level of questions have been observed to be "too detailed" for the grades being assessed.
- The spread of questions across both the term-end exams is not uniform in a couple of grades. Grades VII and IX have comparatively more questions in the yearly exam than the half-yearly exam.
- The range of questions is diverse, ranging between 26 and 85. The largest number of questions are seen for Grade VI.
- Mostly, the questions are constructed in such a manner that they are very easy to answer, encouraging rote learning.
- Several questions are also framed in an ambiguous manner, leading to multiple interpretations being possible while answering.
- In several question papers, all topics have not been covered; questions on grammar-related topics such as भाषा, बोली, लिपि, उपसर्ग, प्रत्यय, तत्सम और तद्भव are missing. Especially in Grade VI, अपठित बोध (reading comprehension) topic is missing entirely in some question papers.
- In several cases, marks allocation for questions is inappropriate and does not conform to the complexity of the questions.
- The structure of the question papers does not conform to the standard templates. Regarding the sequencing of questions in the question papers, "easy" type of questions follow "complex" questions.
- Several question papers have basic errors such as spelling mistakes and missing or incorrect numbering of questions. In some cases, the question paper had only 3 sections instead of the standard 4 sections.
- It is observed that in a school where Hindi is being taught as a second language, question papers are particularly inappropriate both in terms of level and the weightage given to learning outcomes.

The term-end exam papers' questions are largely focused on reading and writing skills with a variation in the spread of number of questions between reading and writing skills across grades. Overall, the question papers have a greater proportion of writing questions than reading questions. **Grade IX** has the highest concentration of writing questions at **78%**, while for other grades, it ranges between **73%** and **75%**. The reading questions hover in the range of **25-28%** across all grades.

Grade IX question papers are of good standard with appropriate coverage of learning outcomes.

Extent of congruence

- Several question papers are grade-appropriate for all the grades.
- The range of questions across grades is fairly uniform, ranging between 26 and 50 in both the term-end exams.
- Some question papers are well-balanced, appropriately covering the required levels of learning outcomes, difficulty level, and levels of Bloom's taxonomy.
- Most of the question papers included questions of medium and easy levels of difficulty. On an average, about 57% medium-level questions are included across grades.
- In general, the question papers covered most of the learning outcomes, focusing on testing the features related to understanding of concepts and their application in an extensive way.

Extent of incongruence

- Some of the yearly exam question papers were seen to have fewer questions with complex difficulty level.
- It is also observed that in some question papers appropriate weightage has not been given to several topics and learning outcomes have not been covered uniformly. *(For instance, 40% of the problems are from Fractions & Decimals in Grade VII with marginal weightage being given to Congruence of Triangles. Topics such as Practical Geometry, Visualising Solids, and Rational Number are missing from Grade VII question papers. In Grade VI, there are no items from topics pertaining to Symmetry and Practical Geometry, and a few important topics such as Algebra and Ratio and Proportion have been completely neglected.)*
- For Grade VIII, consistency in question papers across schools is missing. Further, these questions were mostly knowledge-based.
- In some schools, most of the questions tested LOTS with not enough questions relating to critical thinking, analytical, communication, evaluate, computing, and creative thinking.
- Unusually, in a few schools, the question papers were skewed towards questions with higher difficulty with only around 4% of questions from "easy" level.

In one school, it was observed that the question paper had several questions related to concepts not included in the NCERT textbook. Questions relating to Percentages and Simple Interest were being asked in a Grade VI exam, which are noted as inappropriate for this level.

Extent of conformity

- Some question papers show a healthy balance with a variety of questions from both lower and higher order cognitive skills and of creative nature. A few question papers have included questions that assess higher level of application-based thinking.
- Some question papers also comprise questions related to concepts that have futuristic use.
- Some of the Grade IX question papers have questions assessing different Bloom's taxonomy levels along with a few thought-provoking questions mapped to HOTS.
- There is almost an equal distribution of questions in both the term-end exams for all grades except Grade IX.
- In grades VI and VII, there are about 15% of questions pertaining to HOTS, indicating a satisfactory distribution of questions between HOTS and LOTS.

The distribution of questions among different levels of Bloom's taxonomy improves through Grade VI to Grade IX.

Extent of non-conformity

- For Grade IX, it is observed that the yearly exam has a greater number of questions than the half-yearly exam.
- Most of the question papers are skewed towards lower order cognitive skills. On an average, 94% of the questions are from LOTS in case of grades VIII and IX. In most question papers, there was no scope to capture evidence of creative thinking and analysis.
- Most of the question papers have predominantly text-based questions. There are only a miniscule number of diagram-based questions asked across all the grades. There are no questions pertaining to predicting skills and graphical representation in higher grades.
- Some question papers did not include questions pertaining to the topics of Elements and Compounds and Chemical Formulae. In Grade VII, there are chapters that have not been covered in the question paper. (*These include Ch 16: Water: A Precious Resource, Ch 17: Forests: Our Lifeline, and Ch 18: Wastewater Story*).
- A couple of question papers analysed have a significant number of typographical errors and questions that do not map with any learning outcome.

The learning outcomes as mentioned in the Teachers' Resource for Achieving Learning Outcomes document are of lower order and, therefore, the questions are also predominantly of lower order considering the targeted age group. Also, it is to be noted that the learning objectives stated in the document were rewritten as learning outcomes for the sake of mapping.

THEME 2

Attainment of the desired literacy and numeracy skills—extent to which these are met by the students

Theme 2 aims at understanding how far the students across schools have attained the desired literacy and numeracy skills across grades VI, VII, VIII, and IX and subjects. The predominant criteria that have been considered for this theme are students' scores in English, Hindi, and Mathematics; coverage of learning outcomes, topics, and skills; and students' exposure to attempting passage questions that evoke their understanding of real-life situations.

Broad spectrum observation

There is partial attainment of literacy and numeracy skills across grades.

Extent of attainment

- It is observed that the overall average scores in English and Hindi are almost similar at 72%, showing uniformity in attainment of both.
- In the case of subjects testing literacy, all the question papers include questions pertaining to reading and writing, resulting in the capture of these skills in the overall scores.
- Adequate emphasis on numeracy skills is seen throughout the question papers of Mathematics, resulting in the capture of these skills in the overall scores.
- Some of the question papers have covered most of the learning outcomes pertaining to English, Hindi, and Mathematics across grades, thus providing comprehensive reporting of the attainment of the desired literacy and numeracy skills.

Extent of non-attainment

- Listening and speaking skills in case of literacy subjects have been excluded across grades in the term-end exams, resulting in non-reporting of these skills.
- In both Hindi and English, there is a huge variation in the number of reading and writing questions across grades.
- In English, Hindi, and Mathematics, several learning outcomes have not been covered in several question papers across grades due to which the attainment of these skills is not fully reported.

Listening and speaking skills in case of literacy subjects have been included as part of the periodic tests across grades though they do not figure in the term-end exams.

The scores of the students in English, Hindi, and Mathematics are ranging from **46%** on the lower end to **93%** on the upper end. It has also been observed that the scores of all three subjects do not vary significantly across regions. In comparative terms, Mathematics scores are lower than language scores across all regions with North scoring better in Hindi and South scoring better in English (refer to Fig. 5).

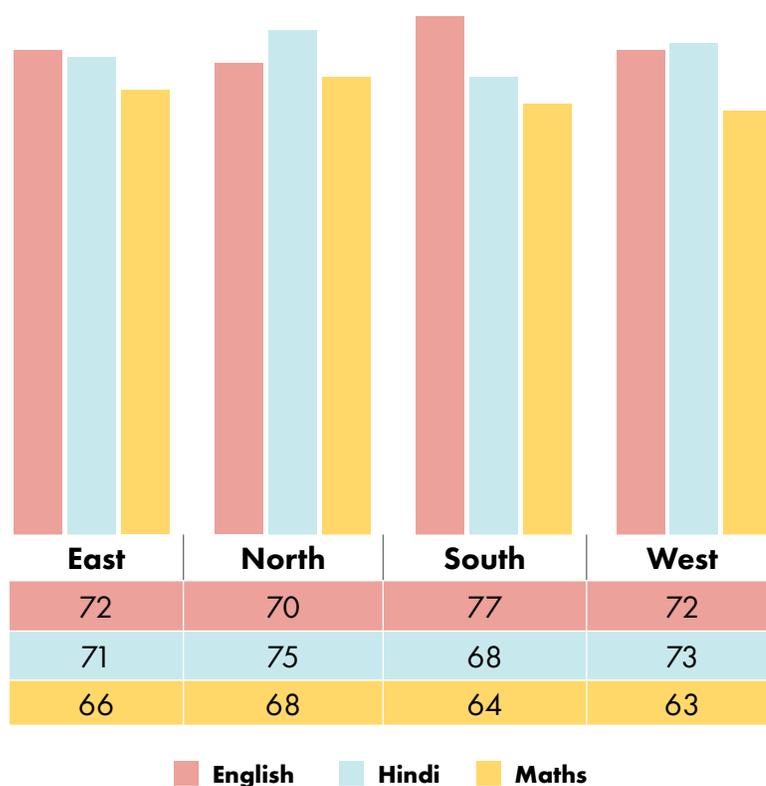


Fig. 5 Region-wise and subject-wise performance

Subject-specific observations

ENGLISH

Extent of match

- The average score across all the grades is around 72%, reflecting a satisfactory performance in terms of attainment of the literacy skills.
- Some of the question papers cover most of the learning outcomes, thus resulting in complete reporting of the attainment of the desired literacy skills.
- Some comprehension questions have well-selected passages that test high-order thinking skills, which are likely to have been included in the reported scores for these students.

Extent of mismatch

- The term-end exam papers do not include items pertaining to listening and speaking. Also, there is variation in the spread between reading and writing questions among the grades. This makes it difficult to determine the extent of attainment of individual skills from the available scores of the students.
- In many question papers, several learning outcomes remain untested, leading to non-availability of scores measuring these outcomes.
- Question papers in general do not include questions pertaining to drawing of inferences and intensive reading, leading to these skills not being tested.

The scores are ranging between 46% and 92% across grades.

HINDI

Extent of match

- The average score is 72%, with Grade VII showing average score as 74%, reflecting satisfactory performance of students in terms of attaining the literacy skills across grades.

The scores uniformly range between 51% and 95% across grades.

Extent of mismatch

- As in English, the term-end exam papers in Hindi also exclude items pertaining to listening and speaking. There is disparity in the spread between reading and writing questions among the grades, making it difficult to determine the extent of attainment of individual skills from the available scores of the students.
- The attainment of about half of the learning outcomes cannot be reported, as these learning outcomes have a coverage of less than 5% across grades in the question papers.
- It has been observed that there are reading comprehension passages across grades that have typographic errors, thereby hampering the ability of the examinees to understand, interpret, and express. This is likely to have impacted the reporting of attainment of these skills.

MATHEMATICS

Extent of congruence

- Extensive emphasis has been laid on assessment of numeracy skills (95% or more) across grades, which have been captured in the scores of the students.
- The average score of Grade VI is the highest with 73% followed by Grade VII at 70%, indicating a satisfactory performance in terms of attaining the numeracy skills across grades.
- Most of the learning outcomes are covered in varying degrees in the question papers of all the grades, resulting in these being reflected, at least to some extent, in the scores of the students.
- A couple of question papers include questions focusing on real-life situations, which is likely to have been included in the reported scores of the students.

Extent of incongruence

- Several question papers have not covered the learning outcomes uniformly and in others appropriate weightage has not been given to topics, leading to difficulty in determining the achievement of these outcomes.
- Largely, question papers across grades lack questions that test mathematical skills. Thus, these skills are not tested and subsequently not reported, making it difficult to check the extent of attainment of the desired skills.
- The average score is steadily falling from 73% in Grade VI to 62% in Grade IX, indicating falling levels of attainment of desired numeracy skills.

The scores of the students range between 43% and 92% for all grades.

THEME 3

Proficiency levels of students when compared with global standards—extent to which students meet the global standards

This theme aims at understanding the proficiency levels of students in comparison to global standards. In the year 2025, students of different regions would be participating in an international standardised achievement test. This makes the Board keenly focused on knowing the standing of its students' learning and skills from a global perspective. The predominant criteria considered here are global competencies and the study of the question papers and student scores in the light of these competencies.

Broad spectrum observation

It is observed that the question papers are limited from a global standards perspective, resulting in lack of feasibility in checking the proficiency levels of the students.

Subject-specific observations

ENGLISH

Most of the questions from the question papers do not conform to global standards. About only 15% of the questions could be mapped to the lower-level of skills/competencies required to perform tasks such as locating information within a text and drawing out the relevant information and literal meaning of words, sentences, and passages.

The questions focusing on the higher-level competencies pertaining to understanding text, making meaningful interpretations by connecting various information given within the text, and applying the developed understanding of the text in different experiential contexts and references are limited. This can be seen as an outcome of the Theme 1 findings that reflect on the quality of question papers.

MATHEMATICS

Most of the questions from the question papers do not conform to the global standards. The presence of questions related to *mathematical literacy* is quite limited. *Mathematical literacy* in this context refers to the skills employed to reason mathematically and apply systematically the various mathematical concepts, formulas, and procedures in different real-life contexts to solve problems.

As observed in Theme 2, about 95% of the question papers focus on basic numeracy skills that involve understanding of the concepts and performing mathematical operations without context. The following figure depicts the grade-wise percentage distribution of questions between mathematical literacy and numeracy skills in the term-end exams. The figure clearly reflects the absence of questions assessing mathematical literacy across grades. However, Grade IX shows a minimal presence of such questions.

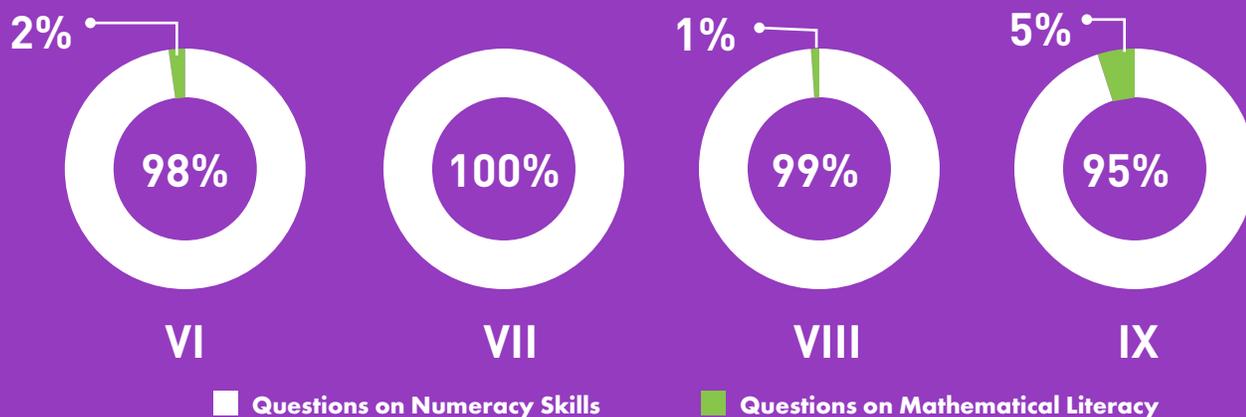


Fig. 6 Grade-wise question distribution—mathematical literacy vs. numeracy skills

SCIENCE

Almost 50% of the questions from the question papers do not conform to the global standards. The remaining questions mostly relate to recalling knowledge of major scientific concepts, procedures, facts, and ideas and explaining the same scientifically.

The number of questions pertaining to scientific analysis and evaluation of data, representation of data, drawing of conclusions, and designing of scientific inquiry to evaluate and test observations is minimum. Questions relating to usage of scientific practices in real-life contexts are largely absent. There are hardly any questions testing epistemic skills across the grades.

THEME 4

Overall achievement of level-appropriate skills and knowledge—extent to which these are met by the students

Theme 4 aims at understanding how far the students of the representative sample set of schools have achieved the overall level-appropriate skills and knowledge across grades and subjects. The predominant criteria that have been considered here are subject-specific total scores of students, achievement of desired learning outcomes, attainment of 21st century skills, and application of knowledge in real-life experiences and situations.

Broad spectrum observation

The achievement of level-appropriate skills and knowledge by the students across grades is not reported adequately, due to unavailability of level-appropriate question papers in term-end exams.

Extent of overall achievement

- The average score of the students is around 70% across grades and subjects, indicating that students' performance is largely satisfactory for the attempted question papers.
- Some of the question papers across grades have thought-provoking and well-designed questions encouraging creative and critical thinking. This is likely to have been reflected in the overall scores of the students who have attempted these questions correctly.
- Some of the question papers cover the required levels of learning outcomes and an appropriate combination of HOTS and LOTS, enabling the reporting of achievement of these skills and knowledge.

It is also observed that the analytical subjects of Mathematics and Science have a stronger correlation of 74% in comparison to linguistic subjects of English and Hindi that show a correlation of 70% (refer to Fig. 7).

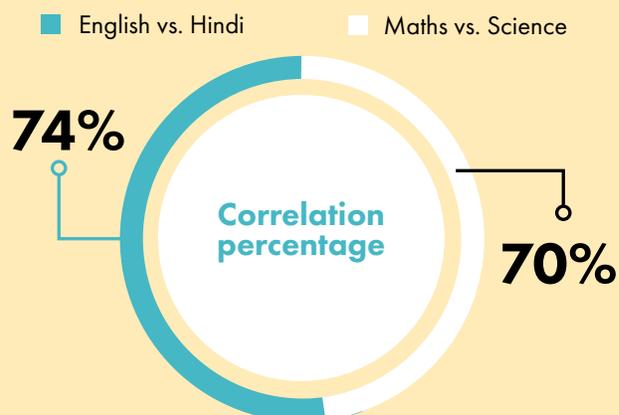


Fig. 7 Correlation percentage—linguistic and analytical subjects

A region-wise analysis of student performance across grades revealed that performances in South across grades seem to be uniform, ranging between 69% and 70%. In comparison to other regions, the South has the highest performance of 70% in Grade IX. In the north region, Grade VI has performed the best, while in East the performance of grades VI and VIII is the highest as well as similar at 71% (refer to Fig. 8).

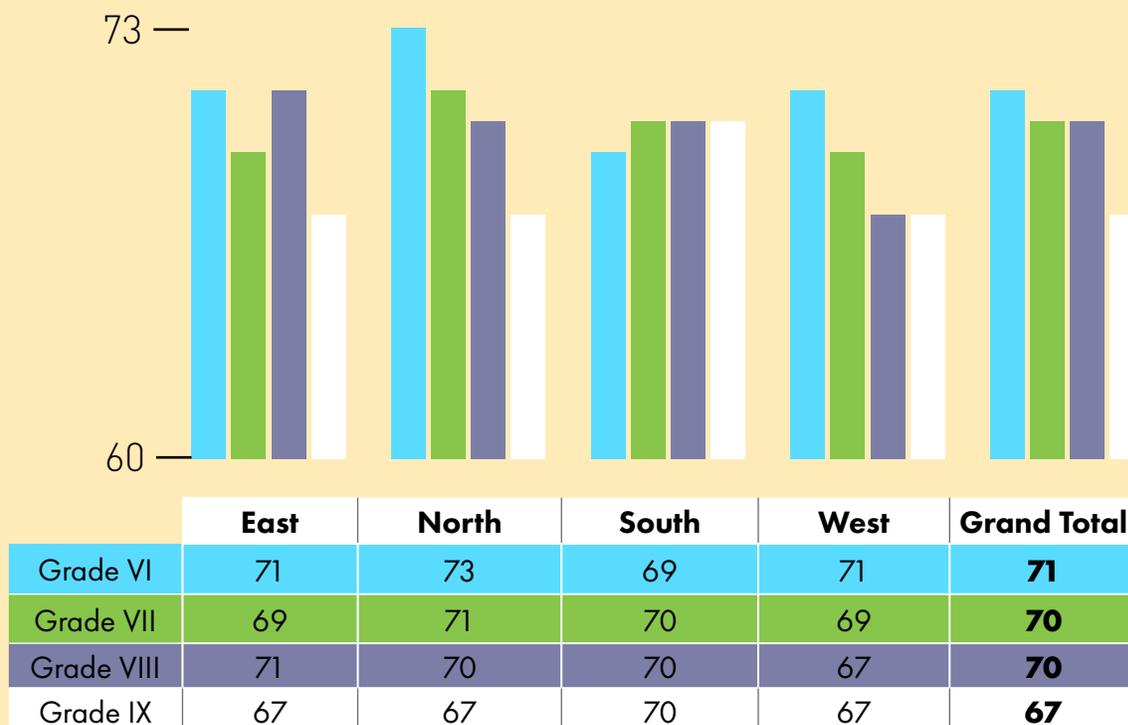


Fig. 8 Region-wise and grade-wise performances (in %)

Extent of overall non-achievement

- The performance of students in general is falling as we move from Grade VI to Grade IX. This may be indicative of progressive gaps in the understanding levels of students, lack of acquiring mastery of grade-appropriate skills and knowledge, inability to cope with increased curriculum load, etc.
- Students' achievements are reflected mostly against only LOTS, due to the presence of minimum number of questions related to HOTS across grades. Further, questions that demand real-life application of concepts have marginal presence in most of the question papers, leading to non-availability of scores related to these skills.
- As several learning outcomes remain untested across grades, students' achievements of level-appropriate skills and knowledge are not adequately reported.

It is observed that Mathematics papers are of a higher-level in comparison to other question papers across grades. This may account for the relatively lower scores obtained by students in Mathematics vis-à-vis other subjects.

Subject-specific observations

ENGLISH

Extent of match

- Students' performance is satisfactory, with average scores being around 72% across grades.
- Some question papers comprise well-selected passages that test higher order thinking skills as well as a few questions based on the experiential learning of the language involving real-life application of concepts. This is likely to have been reflected in the overall scores of students who have attempted these questions correctly.

Extent of mismatch

- There is difficulty in determining the level of achievement of skills such as critical thinking, comparing, interpreting, inferring, analysing, synthesising, and evaluating, as questions related to these skills are either minimal or missing across grades.
- In some cases, the scores of the students cannot be attributed to attainment of level-appropriate skills and knowledge, as these question papers comprise questions that do not fit the purpose of the assessment. The construction of these questions is such that they seem to be more suitable for formative tests than summative exams.
- There is incomplete testing and reporting as term-end exam papers do not include questions pertaining to listening and speaking skills and the opportunities for intensive reading across all the grades are also less.

The average score of Grade IX is 73%, which is higher as compared to grades VI, VII, and VIII.

HINDI

Extent of match

- Students' performance across grades is satisfactory with an overall average score of 72%.
- As several question papers are well-balanced with a good coverage of different levels of Bloom's taxonomy, it allows for testing and reporting of the attainment of learning in a balanced and grade-appropriate manner.

Extent of mismatch

- Across grades, about half of the learning outcomes have a coverage of less than 5%. Thus, these outcomes remain relatively unreported.
- Inadequate percentage of HOTS questions in the question papers has resulted in the lack of testing and reporting of these skills.

At 74%, Grade VII has performed better than the overall average.

MATHEMATICS

Extent of congruence

- The average score of Grade VI is the highest with 73% followed by Grade VII at 70%, indicating satisfactory performance of students in these grades.
- In several question papers, most of the learning outcomes are covered, resulting in reporting of the students' achievement of overall skills and knowledge.
- Overall, the question papers test students' understanding and application in an extensive way with a couple of question papers comprising questions focusing on real-life application of concepts and mathematical literacy. This is likely to have been reflected in the overall scores of students who attempted these questions correctly.
- Some question papers, especially of Grade IX, are balanced with challenging questions, good coverage of levels of Bloom's taxonomy, and appropriate weightage to all topics and learning outcomes. This is likely to have been reflected in the overall scores of students who attempted these questions correctly.

Extent of incongruence

- Performance of students in Mathematics in comparison with other subjects is the lowest.
- The average score is steadily falling from Grade VI to Grade IX, with it being the lowest in Grade IX at 62%. This points towards gaps in understanding and attaining of mastery of grade-appropriate mathematical concepts and skills, attesting to what has already been mentioned in the theme's 'Broad spectrum observation'.
- Students' achievement of overall skills and knowledge is not reported completely as some question papers either lack or have minimal coverage of questions pertaining to HOTS and largely focus on LOTS related to recall and understanding of concepts.

There is significant difference in the average scores of students in lower grades vs. higher grades.

Extent of conformity

- The average score of Grade VI is the highest with 72% followed by grades VII and VIII at 70%, indicating satisfactory performance in terms of the administered papers in these grades.
- In some schools, reporting of the attainment of the overall skills and knowledge through the overall scores of students is comprehensive, due to the presence of a balanced proportion of HOTS and LOTS questions and inclusion of questions encouraging critical and creative thinking. This is especially true in case of grades VI and VII, which have about 15% HOTS questions.
- There are a few questions that test the application of core concepts and scientific practice in real-life situations, resulting in reporting of achievement of these skills across grades.
- Grade IX has some question papers that include questions assessing the different levels of Bloom's taxonomy and a few thought-provoking questions mapped to HOTS. This is likely to have been reflected in the overall scores of the students.

Extent of non-conformity

- In some cases, the format of the question papers is unfriendly for the test takers. This is because in these papers, the questions are not arranged according to Bloom's taxonomy, their structure does not conform to the standard template, and font sizes are inconsistent. This is likely to have adversely impacted the students' scores.
- There is lack of comprehensive testing and reporting of the achievement of grade- and level-appropriate skills. The Theme 1 findings affirm this, as they signify that most question papers are imbued with text-based questions that are largely related to LOTS and lack questions testing predictive skills, graphical representation, and critical and creative thinking, especially in higher grades. Thus, resulting in these skills being unreported.
- Predominantly, the question papers lack questions that have connection to real-life situations, as around 90% of questions are theoretical in nature. This has led to non-reporting of achievement of these skills.
- A couple of question papers have questions that do not map to any learning outcome, undermining the reporting of the achievement of the overall skills.

Across the schools, the scores range between 48% on the lower end and 93% on the upper end.

Recommendations

For the current assessment system, to be able to implement the NEP's (2020) vision of *assessments being tools of learning* and become more effective, having a deeper learner-centric approach is essential. The following metrics are approaches that can be administered to not only further the attainment of level-appropriate skills and knowledge by students but also effect NEP-recommended changes.

01 | Developing a comprehensive assessment framework

- Holistic
- Competency-based
- 21st century skills
- Subject and grade specific

02 | Framing a measurable blueprint for test construction

- Adheres to assessment framework
- Defines test specifications
- Is flexible
- Ensures consistency in question papers

03 | Ensuring rigour and quality in assessments

- Adhere to the specified blueprint
- Follow assessment principles of validity, reliability, fairness, usability, inclusivity, and relevance
- Comply with quality assurance

04 | Specifying evaluation measures

- Broader goals at grade level
- Grade-specific performance expectations/tasks
- Achievement targets, scales and achievement descriptors
- Evidence-based evaluation

05 | Building assessment capacity

- Assessment framework and blueprint implementation
- Item development
- Reporting
- Data management, analysis, and interpretation

DEVELOPING A COMPREHENSIVE ASSESSMENT FRAMEWORK

One of the key metrics is developing a comprehensive **assessment framework** that acts as a backbone to building assessments based on the expected standards and learning progressions to measure the achievement of educational goals.

The implementation of the Board's core objectives across its schools can be operationalised by developing the assessment framework with a defined set of **assessment objectives** addressing the grade- and subject-specific needs along with focus on foundational needs and encompassing different purposes of assessments (refer to Fig. 9).

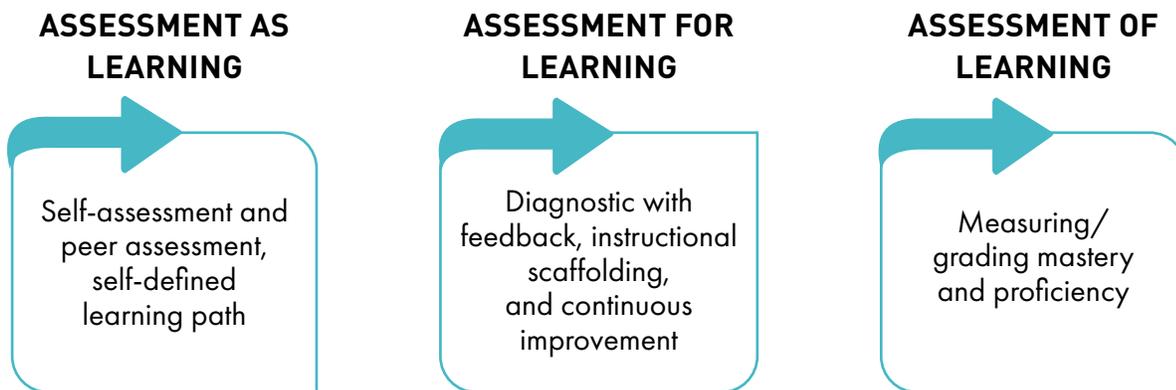


Fig. 9 Different purposes of assessments

To enable the current assessment system to move away from rote learning and towards meaningful learning, the framework needs to be holistic and competency-based, and include both the **scholastic** and **co-scholastic** aspects of students' learning journey. Such a framework would allow for overall development of learners through socio-emotional learning as advocated in NEP 2020.

To create future-ready citizens who can succeed in the real world and have a global standing, understanding of academic content must be complemented by intertwining **21st century skills**—communication, collaboration, environmental awareness, financial and economic literacy, Information and Communication Technology (ICT) literacy—within the subjects.

Accordingly, the assessment framework needs to cater to different dimensions related to subjects, cognition, metacognition, and attitude with focus on global competencies.

The design of the framework is usually indicative of the natural thinking progression of individuals and the evolution of understanding of content knowledge, clearly defining the standard of performance expected at each grade so that the assessment tasks can provide ample opportunities for demonstration of mastery of level-specific skills and knowledge.

FRAMING A MEASURABLE BLUEPRINT FOR TEST CONSTRUCTION

A measurable blueprint helps in assessing the knowledge, competencies, and skills adequately.

For this, the blueprint needs to adhere to the assessment framework and clearly state the test specifications as per the purpose of the test, and guide creation of balanced grade-appropriate question papers that include fit-for-purpose questions with a suitable distribution of topics and complexity levels. Thus, having a blueprint ensures that the achievement of the required fundamental skills is also being measured.

Such a blueprint is typically **flexible** to incorporate requirements catering to regional and cultural contexts as recommended in NEP 2020; at the same time, providing a **level of consistency** in the question papers to ensure equity and inclusivity.

The blueprint usually defines the number of items to be included in the test, the length and breadth of the subject-matter to be covered, appropriate percentage of items for testing HOTS, level of difficulty, and questions pertaining to knowledge, demonstration in real-life situations, and communication of concepts. It may also include details specific to the purpose of the test, test time, weightage for topics, skills and competencies to be assessed, and item formats.

Constructing items and designing question papers that adhere to the blueprint will ensure **satisfactory measurement** of the level-appropriate competencies, skills, and knowledge, thus providing **a clear picture of the attainment** of these skills as well.

A differentiated blueprint can be prepared for different types and purposes of assessments. For example, formative tests are used as a means to bring students towards the standards that need to be achieved in the summative tests. Such nuanced differentiation in the purposes of different types of assessments creates the need of having a separate blueprint that specifically caters to the purpose and type of assessment.

RECOMMENDATION 03

ENSURING RIGOUR AND QUALITY IN ASSESSMENTS

As the NEP 2020 states, the eventual purpose of assessments is to enable improvement in student learning. The assessment results are a means of getting useful information regarding what a student has learned and achieved in terms of the desired skills, knowledge, and learning outcomes.

For assessments to be able to provide this information with accuracy to act as a guide for taking remedial actions, it requires adherence to the specified blueprint and the assessment principles of **validity, reliability, fairness, usability, inclusivity, and relevance**.

A well-defined set of guidelines would provide the necessary actions for conforming to the **framework** and **blueprint**. A credible process on **quality assurance** of tests will help in ensuring acceptable test standards, including **accessibility** and **readability** of assessments.

RECOMMENDATION 04

SPECIFYING EVALUATION MEASURES

One of the first steps towards establishing evaluation techniques is to define the **broader goals at grade level that are subject-specific** and align the assessments to these goals. This is to be followed by defining the **performance expectations/tasks** for each grade and subject. These tasks or expectations need to be aligned to the learning outcomes and further act as a summary of the knowledge and skills a student should attain **mastery** of at a particular grade level.

This would allow for the realisation of NEP 2020's recommendation related to improving evaluation standards.

It is imperative to use a set of **criteria/rubrics** that include achievement targets, scales, and descriptions against the achievements to explicitly define the assessment of a specific performance. Illustrating what performance would mean or look like if a student exceeds, meets, approaches or is below the expectations is represented through a **Spectrum of Criteria** that tracks the presence/absence of evidence of performance of the students. Rubrics help in identifying the competencies attained through proficiency judgements.

Such rubrics would help in providing appropriate weightage to a specific topic and/or competency as per its importance for that particular grade, subject, and level. It would

also enable the creation of a suitable marking scheme that allocates scores as per the significance of the topic/competency/skill being assessed.

Precise rubrics lay the ground for **evidence-based evaluation** by providing definitive evaluation criteria mapping to the level of achievements, thus reducing grading time and increasing objectivity while grading.

These evaluation techniques are accompanied with inputs related to progress and remediation. It is suggested that the remedial measures be built on the pillars of **instructional scaffolding** and **guided learning** to enable the students to gradually move up in the learning progression scale.

While instructional scaffolding will provide the structure and support needed to facilitate development and mastery of a concept, guided learning will provide the focused intervention, interaction, and collaboration required to help students acquire the competency to perform a task independently.

RECOMMENDATION 05

BUILDING ASSESSMENT CAPACITY

Assessments are important for educators, as they can provide them with the relevant qualitative and comparative information to monitor and evaluate students' progress as well as review and improve the teaching-learning process.

It is, therefore, appropriate to support teachers in using and implementing assessment framework and blueprint and applying strategies for achieving effectiveness and robustness in assessing the students.

This is in-line with NEP 2020's focus on empowering teachers through continuous opportunities of training in the latest assessment strategies and pedagogies.

A clear set of guidelines, trainings, and support documents with FAQs related to **item development, reporting, data management, and analysis and interpretation of students' scores** would support teachers and/or item developers in facilitating the achievement of the desired learning outcomes through evidence-based decision-making.

With a focus on the benefits of competency-based assessments, global competencies, and importance of foundational skills and knowledge, training needs to direct the thinking of teachers towards creating tasks that enable effective evaluation of the outcomes of learning.

A bank of sample assessments will provide the necessary source of guidance for teachers and administrators to develop the expected types of assessments.



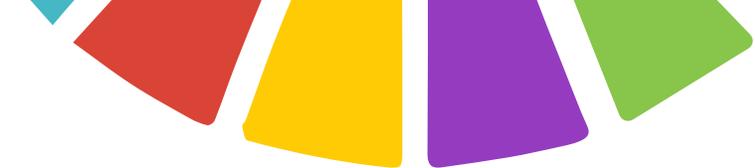
Conclusion

The study of the modified system of assessment and examination that was introduced by CBSE from Academic Year 2017-18 onwards has found that the question papers administered during the term-end exams align only in part with the standards set by the Board.

The approach taken in this research project was beneficial in gathering empirical evidence that enhanced the understanding of the assessment patterns being followed across subjects—English, Hindi, Mathematics, and Science and grades VI-IX. It highlighted the assessment tasks that were robust examples of assessing skills such as creativity and critical thinking through application in real-life contexts.

However, it also brought forth the degree of incompatibility of the current assessment patterns with the NEP 2020's vision of having more “regular, formative, and competency-based assessments” that test higher-order thinking skills and are used primarily as “assessment for learning”. This variance between the current and future assessment patterns also undermines the multidisciplinary and cross-cutting concepts learning that NEP 2020 is aiming towards.

The study highlights the need for designing an assessment structure that assesses the deep learning of the 21st century skills such as critical thinking, problem solving, collaboration, citizenship, creativity, etc. to not only implement the aims of NEP 2020



but to also create future-ready students who are equipped with the right skills to work and study anywhere in the world. Such a structure would also further the Board's purpose of creating a learning environment that nurtures holistic development of students.

This mammoth task can be broken down into smaller steps as suggested in the recommendations above. The Board, in collaboration with assessment specialists, can move towards transforming its assessment system into a system that provides opportunities for demonstration of skills, aptitude, and cognitive capabilities.

This assessment system should allow teachers to observe the skills in action and familiarise themselves with them while assessing the students. To develop such a system, data collection needs to continue on a larger scale so that evidence of the behaviours being demonstrated and how these behaviours link together to form levels of development can be captured.

Further data collection and analysis on the following lines can allow for deeper investigations, leading to an assessment culture that focuses on demonstration of skills rather than mere recall of content knowledge.

- How do individual students perform across the different skills when measured simultaneously in the same context?
- How can skills be interwoven to measure cross-cutting concepts, application of knowledge, and experiential learning?
- How can this help us understand how to teach them?
- How do the skills develop from Grade VI to Grade IX?
- What does the rate of growth look like for these skills?
- How, if at all, do the skills transfer across different learning areas?

Appendix

Items – Examples and Non-Examples

Following are some items from the actual data that showcase examples and non-examples of good items.

ENGLISH

Example- Grade 8

You are Sunny. You are happy that your childhood friend helped you in getting your son released from a false case of dacoity. Write a letter to your friend thanking him for his timely help.

Non-example- Grade 7

Your papa took you to a local orphanage to distribute sweets and stationery goods among children there. You felt deeply moved seeing the children without parents and also felt a lot of satisfaction at helping them. Express your feelings in the form of a diary entry.

(Hints: eye-opening experience, committee of that orphanage, don't get love of their parents, aware about the plight of countless children, met with shriek of joy, felt a great satisfaction, donate for their welfare, come back with new lesson)

SCIENCE

Example- Grade 8

Akanksha has made two paper cups by folding a sheet of paper. She half fills one of the cups with water and keeps the other empty. She puts those two paper cups above the flame carefully. With reference to the above activity, answer the following questions.

- Which zone of the flame is best for heating?
- What happens to the empty paper cup when it is heated?
- What happens to the paper cup with water and why?
- What will be the change in temperature of water?

Non-example- Grade 9

Define Power.

MATHEMATICS

Example- Grade 6

Mary had a piece of ribbon that was 60cm long. She used $\frac{3}{5}$ th of it to make bows and $\frac{2}{3}$ rd of the remaining to wrap a box.

- What was the length of the ribbon used to make the bows?
- How much of the ribbon was left with her?

Non-example- Grade 8

A parallelogram with all sides equal is called a

- Trapezium
- Rectangle
- Rhombus
- None of these

HINDI

Example- Grade 8

(ख) दिए गए वाक्य में उचित विराम चिह्न लगाइए-

- तुलसीदास जी ने रामचरितमानस लिखी थी
- मैंने आदरपूर्वक पूछा, देवी क्या आज्ञा है

Non-example- Grade 6

4. कोष्ठक में दिए गए निर्देशानुसार उत्तर दीजिए-

- कपड़ा, घर (एक-एक पर्यायवाची शब्द लिखिए)
- आशा, उपकार (विलोम शब्द लिखिए)

Acknowledgements

This research study would not have been possible without the exceptional support of the following people from the Central Board of Secondary Education (CBSE) and Oxford University Press (OUP).

Their knowledge and guidance at every step of the way has helped bring this study to fruition.

CBSE

Manoj Ahuja, *Chairman*

Dr Biswajit Saha, *Director, Training & Skill Education*

Dr Joseph Emmanuel, *Director, Academics*

Dr Sandeep Kumar Jain, *Joint Secretary, Training*

Dr (Mrs) Sweta Singh, *Joint Secretary, Academics*

Oxford University Press

Fathima Dada, *Managing Director, Education*

Claudia Bickford-Smith, *Director of Education Services*

Sumanta Datta, *Managing Director, India*

Dave Leach, *Director, Global Assessments and Project Director, PISA Science Framework Development*

Dr Jonathan Osborne, *Kamalchari Professor of Science Education, Stanford Graduate School of Education and Chair, OECD PISA Science Expert Group, Advisor on behalf of OUP*

Yash Mehta, *Executive Director, Sales, India*

Paras Bansal, *Product Director, India*

S. Uma, *Head, Global Assessments*

Priyadarshini Bhattacharya, *Head, Market Services*

Akansha Yadav, *Assessment Manager*

