

Student Learning Study

Status of Student Learning across 18 States of India in Urban and Rural schools



We acknowledge the support given by the states for conducting the study.

The study is conceived and executed by:



Educational Initiatives Pvt. Ltd.
www.ei-india.com

The study is supported by:



ACKNOWLEDGEMENT



From the Authors:

Student Learning Study would not have been possible without the help and support of a large number of people.

We would like to specifically thank the State Governments for giving permissions to freely conduct the study in the schools for classes 4, 6 and 8.

We also thank the suggestions received from Ministry of Human Resource Development, India; NCERT, EdCIL and Technical Consultant Group of SSA on different aspects of the study design.

We would like to acknowledge the support from Prof. Arun C. Mehta, and the National University of Educational Planning and Administration (NUEPA) for providing the raw student enrolment data of the different states for enabling sampling. We also thank the state officials who supplemented this data wherever required.

We thank Prof. Bimal Roy and Prof. Saurabh Ghosh, Indian Statistical Institute, Kolkata for their support and guidance on the sampling plan for the study.

We thank Dr. Daniel Grath, Dr. Eugene Owen, Dr. Steve Norman from the National Centre of Educational Statistics, Washington and Dr. Keith Rust, sampling referee for PISA for guidance and suggestions on the sampling based on their experience from large scale assessments such as the NAEP, TIMSS and PISA. We would like to acknowledge the suggestions provided by Dr. Amita Chudgar, Michigan State University for collection of background variables.

We would like to express our gratitude to the Google Org team for their continued support throughout the study.

Lastly, but in no way least, we were able to conduct the study smoothly because of the support and encouragement provided by the students and teachers to the field team who worked tirelessly to roll out the assessment under standardised conditions.

MR. SRIDHAR RAJAGOPALAN
(Managing Director)

MS. VYJAYANTHI SANKAR
(Vice President – Large Scale Assessments)

THE TEAM



Google Org

Ms. Linda Segre
Ms. Salimah Samji
Mr. Scott Coleman
Ms. Juliette Gimon

Educational Initiatives

Mr. Sridhar Rajagopalan, Managing Director
Mr. Venkat Krishnan N, Director
Ms. Vyjayanthi Sankar, Vice President – Large Scale Assessments
Mr. Ravi Manoj, Project Manager
Mr. Ashok Mutum, Zonal Manager, North-East
Ms. Gauri Sharma, Zonal Manager, North West
Ms. Kanupriya Misra, Zonal Manager, North
Ms. Krishna Narayan, Zonal Manager, Central
Mr. Madhu Gottumukkala, Zonal Manager, South
Mr. Raghav Rohatgi, Project Manager, ASSL
Ms. Archana Dwivedi, Research Fellow
Ms. Jayanthi Somasekhar, Research Fellow
Ms. Sailaja Ravi, Junior Research Fellow
Mr. Alex Rios, Data Analyst
Ms. Asmi Arul, Educational Fellow
Mr. Alok Mishra, Educational Specialist
Mr. Ashtu Killimangalam, Educational Specialist
Ms. Ekta Hattangady, Educational Specialist
Ms. Manju Bhatnagar, Educational Specialist
Mr. Maulik Shah, Educational Specialist
Ms. Nandini Shah, Educational Specialist
Mr. Shashank Bansal, Educational Specialist
Ms. Swara Pandya, Educational Specialist
Ms. Neeti Bhagchandani, Sr. Executive - Training
Ms. Anar Shukla, Manager - Online Marketing
Mr. Rohit Sivakumar, Sr. Manager - IT & Systems
Mr. Arpit Metaliya, Software Specialist
Mr. Chaitan Sukhadia, Jr. Software Engineer
Mr. Devpal Shah, Senior Graphic Designer
Mr. Rajendra Soyantar, Graphic Designer
Ms. N.V.N.S.S.Durga, Junior Graphic Designer
Ms. Sowmya, Junior Graphic Designer
Mr. Ketan Behrawala, Asst. Manager - Accounts
Ms. Sheel Shastri, Asst. Manager – Finance
Ms. Rupande Shah, Asst. Accountant
Mr. Shivang Dave, Sr. Manager - Logistics
Ms. Jignasha Mistry, Logistics Executive
Mr. Ravikant, Office Assistant

Educational Initiatives – Field Team

Mr. Anand Paul, District Coordinator, West Godavari
Mr. Anindya Basu, State Coordinator, West Bengal
Ms. Anuradha Sinha, District Coordinator, Sultanpur
Mr. Arjun Kumar, District Coordinator, Mahaboobnagar
Ms. Asha Molker, District Coordinator, Gulbarga
Mr. Ashish Sharma, District Coordinator, Jaipur, Jodhpur, Ludhiana and Tehri
Ms. Ashmeet Kaur, District Coordinator, Chandigarh

Mr. Bhaskaran, District Coordinator, Dharmapuri
 Mr. Bijendra Singh, District Coordinator, Saharsa, Bihar
 Mr. Binod Singh, District Coordinator, Patna, Bihar
 Mr. Devesahayam, District Coordinator, Pudukottai
 Mr. Dhiraj Kumar, District Coordinator, Saran, Bihar
 Mr. Digambar Tulle and Ms. Rajshree, District Coordinators, Nashik District
 Mr. Filman Oraon, District Coordinator, Chatra, Jharkhand
 Mr. Kishore Martha, District Coordinator, Koraput
 Ms. Lakshmi Prasad, District Coordinator, Chennai
 Mr. Lalit Chilhate, District Coordinator, Bhopal
 Mr. Laxman, District Coordinator, Waynad
 Mr. Monesh Gawande, District Coordinator, Betul
 Mr. Parashuramulu, District Coordinator, Hyderabad
 Mr. Piyush Mishra, State Co-ordinator, Madhya Pradesh
 Ms. Pooja Awasthy, District Coordinator, Hoshangabad
 Mr. Pradeep Korde, District Coordinator, Yavatmal, Maharashtra
 Mr. Prashant Pawar, District Coordinator, Nagpur, Maharashtra
 Mr. Santosh Kumar Roy, District Coordinator, Araria, Bihar
 Mr. Santosh Pradhan, State Co-ordinator, Orissa
 Ms. Sarala Kumari, District Coordinator, Trivandrum
 Ms. Sarojini, District Coordinator, Kanya Kumari
 Mr. Sayed Taheer, District Coordinator, Delhi and Gurgaon
 Mr. Sazidur Rahman, District Coordinator, Dhubri, Assam
 Mr. Shafin Ali, District Coordinator, Ranchi, Jharkhand
 Mr. Shashi Krishna, State Coordinator, Bihar
 Mr. Shrey Singh, District Coordinator Delhi
 Mr. Shyam, District Coordinator, Dungarpur, Tonk, Sangrur and Dehradun
 Mr. Srinivasulu, District Coordinator, Kadappa
 Mr. Suryavanshi, District Coordinator, Nanded, Maharashtra
 Mr. Ved Prakash, District Coordinator, Lucknow and Sitapur
 Mr. Vikas Bhale, District Coordinator, Rajnandgaon
 Mr. Virander Singh, District Coordinator, Jammu
 Ms. Vishakha Bhale, District Coordinator, Panna and Raipur
 Mr. Yax Dave, District Coordinator, Surendranagar

External Experts

Ms. Aparna Sridhar, Gujarati
 Dr. A.K.Basu, Bengali
 Ms. Bhakti, Hindi
 Mr. Bhaskar J. Sharma, Assamese
 Ms. Chaya Devi, Telugu
 Mr. Deepak, Hindi
 Dr. S.N. Gananath, Kannada
 Ms. Gayatri Sriram, Tamil
 Ms. Jonaki Bhattacharya, Bengali
 Mr. Jayesh Adhyaru, Gujarati
 Ms. Karkuzhali Sreedhar, Tamil
 Ms. Krupa Gandhi, Hindi
 Mr. Madhusudhanan, Malayalam
 Ms. Manjari Gautam, Hindi
 Ms. Mousumi Bora, Assamese
 Mr. Nayan Bora, Assamese
 Ms. Neeta Thaker, Gujarati

Ms. Padma G., Malayalam
 Ms. Pritam P. Goswami, Assamese
 Mr. Rajagopalan, Tamil
 Dr. K. S. Rajyashree, Marathi
 Mr. Rakesh Sharma, Punjabi
 Ms. Ramadevi, Telugu
 Dr. Ranjit S. Rangila, Punjabi
 Mr. Saheed Ali, Malayalam
 Dr. Sam Mohan, Tamil
 Dr. Sarojini Devi, Malayalam
 Ms. Subhra Sahoo, Oriya
 Ms. Suchita Sane, Marathi
 Ms. Sumangala, Kannada
 Ms. Swarnali Chaudhary, Bengali
 Ms. Vasantha, Telugu

INDEX



EXECUTIVE SUMMARY

9

1. DESIGN OF THE STUDY

1.1.	Need for the Study	17
1.2.	About the Study	17
1.3.	Salient Features of the Study	19
1.4.	Study Coverage	21
1.5.	Question Paper Design	21
1.6.	Background Questionnaires	26
1.7.	Video Study	26

2. OVERALL LEARNINGS

2.1.	Main Findings	27
2.2.	Subject Wise Misconception and Common Errors	41
2.3.	Learning across Classes	47
2.4.	Free Response Items that required Writing	51
2.5.	Language Oral Reading Test	55
2.6.	Background Factors	55

3. COMPARATIVE FINDINGS

3.1.	Performance of Different States	64
3.2.	Comparative Performance of Boys and Girls	66
3.3.	Comparison across Urban and Rural India	66
3.4.	Comparison with National and International Performance	67

4. SLS 2009 BENCHMARKS

4.1.	What are Benchmarks	71
4.2.	Benchmark Descriptions	73
4.3.	How Different States achieve the Benchmarks	75

5. RECOMMENDATIONS

5.1.	Overall Recommendation and Policy Suggestions	80
5.2.	Limitations	85

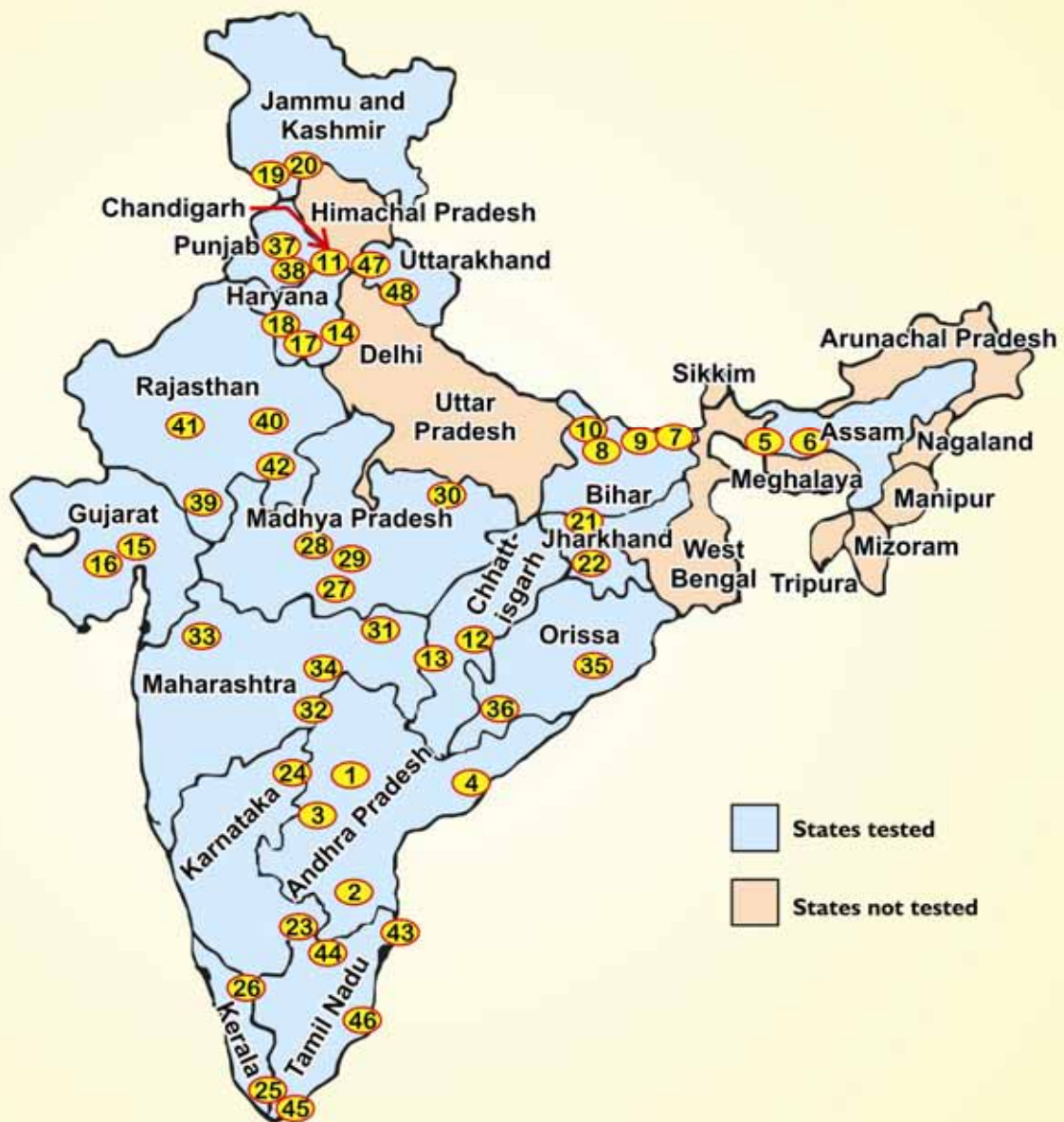
6. SAMPLE DESIGN

6.1.	Purpose	87
6.2.	Steps Followed for finalising the Sampling Design	87
6.3.	Collection of Enrolment Data	87
6.4.	Methods and Details of Sample Design	87

7. TEST DEVELOPMENT

7.1.	Built on the Earlier Municipal School Benchmarking (MSB) Study	89
7.2.	Textbooks Analysis	90
7.3.	Basis for Test Design	91
7.4.	Process of Test Development	91

8. TEST ADMINISTRATION	
8.1. Permissions	93
8.2. Collection of available Enrolment Data and Statistics in all States	93
8.3. Recruitment	93
8.4. Master Training Workshops	94
8.5. Evaluators Workshops	94
8.6. Standardisation Processes for Field Operations	95
9. REPORT AND ANALYSIS	
9.1. Methods of Data Handling	98
9.2. Analysis Methods	98
BIBLIOGRAPHY	99
APPENDIX	
Appendix A: List of Competencies	101
Appendix B: Overall Summary Statistics of All Papers	103
Appendix C: SLS 2009 Benchmarks	104
Appendix D: State Level Statistics for All papers	141
Appendix E: Multiple Comparisons of Average Achievement for States	144
Appendix F: State Wise Rural and Urban Statistics for All Papers	147
Appendix G: IRT Parameters for All Questions	150
Appendix H: Matrix Plot for All Papers	156
Appendix I: Supporting Information from Analysis of Background Factors	157
Appendix J: State Wise - Oral Test Overall Performance	163
Appendix K: Permission Letters	164
Appendix L: Frequently Asked Questions about the Project	170
Appendix M: Master Trainer Manual	173
Appendix N: Evaluator Manual	183
Appendix O: District Coordinator Manual	189
Appendix P: Recruitment Poster, Test Paper and Scoring Sheet	201
Appendix Q: Master Trainer and Evaluator Feedback Form	204
Appendix R: Evaluator Training Attendance Sheet	205
Appendix S: Enrolment Feedback Form and FAQ	206
Appendix T: Evaluator Certificate and College letter	209
Appendix U: School Information Schedule	210
Appendix V: Student Questionnaire	214
Appendix W: Top Sheet and Scorecard	215
Appendix X: Question Papers	216



LEGEND

Andhra Pradesh	Delhi	Kerala	Punjab
1 Hyderabad	14 Delhi	25 Thiruvananthapuram	37 Ludhiana
2 Kadapa	Gujarat	26 Waynad	38 Sangrur
3 Mahabubnagar	15 Ahmedabad	Madhya Pradesh	Rajasthan
4 West Godavari	16 Surendranagar	27 Betul	39 Dungarpur
Assam	Haryana	28 Bhopal	40 Jaipur
5 Dhubri	17 Gurgaon	29 Hoshangabad	41 Jodhpur
6 Kamrup	18 Rewari	30 Panna	42 Tonk
Bihar	Jammu and Kashmir	Maharashtra	Tamil Nadu
7 Araria	19 Jammu	31 Nagpur	43 Chennai
8 Patna	20 Kathua	32 Nanded	44 Dharmapuri
9 Saharsa	Jharkhand	33 Nashik	45 Kanniyakumari
10 Saran	21 Chatra	34 Yavatmal	46 Pudokottai
Chandigarh	22 Ranchi	Orissa	Uttarakhand
11 Chandigarh	Karnataka	35 Khordha	47 Dehradun
Chhattisgarh	23 Bangalore	36 Koraput	48 Tehri Garhwal
12 Raipur	24 Gulbarga		
13 Rajnandgaon			

The Student Learning Study conducted by Educational Initiatives covered 18 major states of India. Between 2 and 4 districts were sampled in each state. Over 100,000 students of classes 4, 6 and 8 from 2,000 schools took tests in Maths and Language in 13 different mediums of instructions.

EXECUTIVE SUMMARY



How do we measure the quality of school education? World over, governments and citizens alike seem to agree that good *quality* education (rather than mere school attendance) is powerfully related to individual outcomes in the labour market, enhancing economic growth of a nation and reducing inequality in society (Hanushek and Woessmann 2007; Vegas and Petrow, 2008; World Bank, 2007).

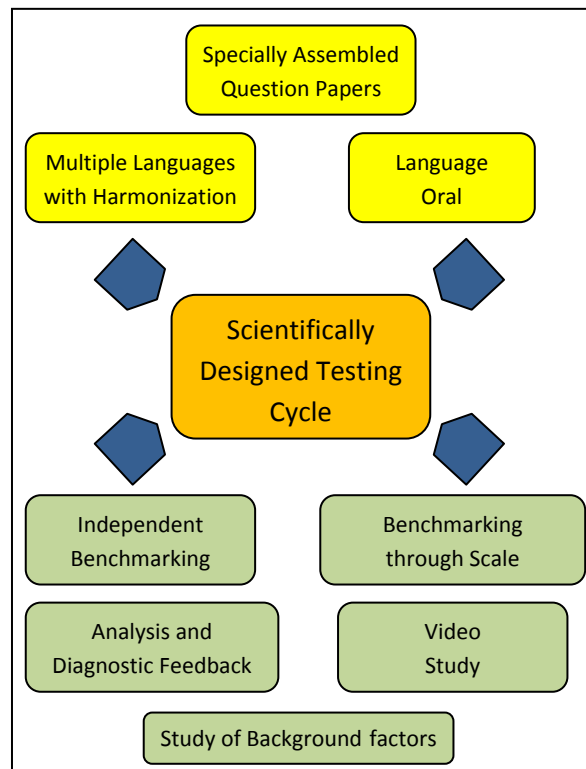
Student Learning Study (SLS) is a benchmarking study of student learning conducted by Educational Initiatives (EI), with financial support from Google.org, USA. The study has been carried out in 48 districts in 18 states and 1 Union territory of India. About 1.6 lac students studying in classes 4, 6 and 8 in 2399 selected government schools were sampled (101643 students actually took the tests – the others were absent on the day of testing) and tested in Language and Maths through common test papers in 13 language versions. The study also collected background information from the students, teachers, head teachers and schools to detect relationships between these factors and student learning.

Other studies like Pratham's ASER (ASER 2005-09) and NCERT's achievement surveys (NAS, 2008) have revealed low student learning levels. The current study is different in 3 important ways – the test design is more sophisticated and checks for student understanding (not just procedural learning). It is independently conducted by a single external agency (EI) and the analysis includes advanced methods like Scale Anchoring. We believe that detailed data helps to devise remedies and solutions in addition to highlighting the problem.

Features	EI SLS	ASER	NCERT Study
Class-wise tests	Class wise tests	Common test	Class wise tests
Single Agency	Facilitated and carried out by Educational Initiatives	Facilitated by Pratham and carried out by local organizations or institutions in different districts	Facilitated by NCERT and carried out by the different states
Tests in School	In classrooms	In homes	In classrooms
Classes tested	Classes 4,6,8	Ages 6 -14	Classes 3,5,7
Coverage	Urban and Rural	Rural	Urban and Rural
Testing Tools	Full length included all competencies; multiple choice and free response questions involving written answers; additionally oral reading tested	Short test with items in oral reading, subtraction, telling time and currency tasks	Full length, entirely multiple choice questions, written responses limited to few in Language paper
Analysis	Achievement data of states with detailed diagnostic feedback capturing learning gaps and misconceptions	Achievement data of states	Achievement data of states
Scale Anchoring	Scale Anchoring using modern Item Response Theory (IRT) to release benchmarks on what students know and do at different ability levels	No Scale Anchoring	No Scale Anchoring

About the Study: The study is probably the biggest of its kind in one country. It was conducted between January and September 2009 in 18 states across India. The main features of the study are:

- Scientifically developed full-length class-specific tests in language and maths designed to test not just recall or procedure but understanding and application.
- Equivalent test forms developed in 13 languages for testing in classes 4, 6 and 8 in the different states.
- Test conducted with permission from state governments. 3 states (HP, UP and WB) could not be covered due to lack or delay in receipt of permission.
- Apart from written component, oral reading test conducted on a sample of students.
- Detailed analysis and diagnostic feedback including benchmarks through Scale Anchoring. A video study was conducted in 3 states to analyse student misconceptions.



MAIN FINDINGS

- 1. Learning levels are extremely low.** In the lower classes, a fair amount of ‘rote-based’ or ‘procedural’ learning is evident in very basic numeracy skills such as number sequencing, operations of whole numbers involving 1 or 2 digits; naming of numbers; reading clock time, understanding currency, etc. Even among procedural questions; students are able to comparatively handle only ‘straightforward’¹ questions that are closer to what one would practise from a typical textbook and not when they are slightly atypical. In higher classes, students are falling behind in all learning, even procedural. Responses to some questions suggest that students are probably coping through learning happening outside the class. For example, in class 6, while more students could add the fractions $2\frac{1}{2}$ and $1\frac{1}{2}$ in a word problem using a real life context, fewer of them could add the same when asked as a straightforward addition question as ‘ $2\frac{1}{2} + 1\frac{1}{2} = \underline{\hspace{1cm}}$ ’.
- 2. Learning taking place is not ‘Learning with understanding’ and a number of misconceptions exist among students on the concepts learnt.** The learning that is happening seems to be procedural or rote-based and not one of ‘Learning with Understanding’ as students find it difficult to answer questions that require a deeper understanding of the concept.

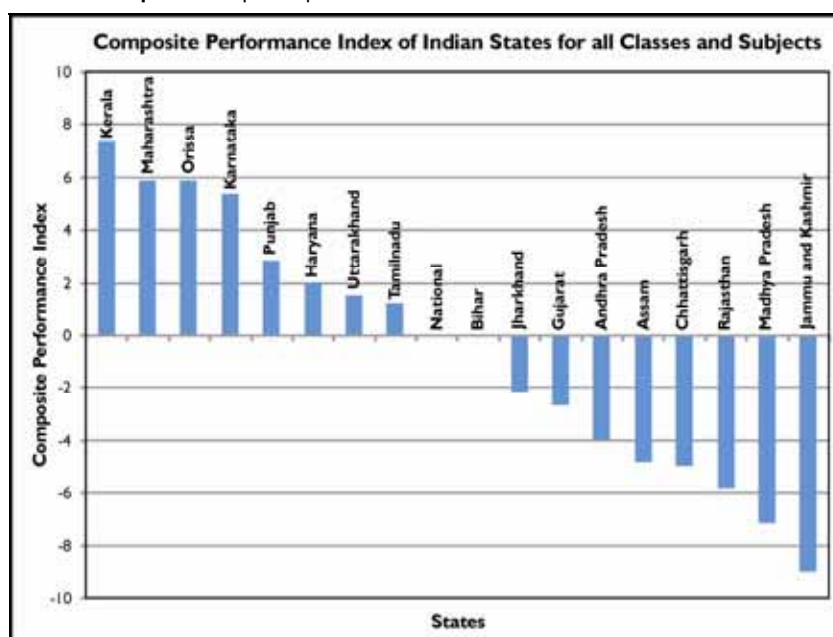
Class 4	Maths
Write the answer. 43 × 2 _____ _____	Fill in the appropriate number in the box. $3 \times \square = 3 + 3 + 3 + 3$
Sample Question 25: This is a procedural question that checks for the process of multiplication. 67.1% of students answered this correctly.	Sample Question 26: This is a conceptual question that checks whether the student is able to link multiplication with repeated addition of a number. 30.4% of students answered this correctly.

¹ A straightforward question is one which has a ‘form’ as it appears in the textbook. For example, a question that asks what 4 tens and 3 ones is considered straightforward; whereas one that asks what is 5 ones and 4 tens – in which the order is reversed and which tests whether the child has understood the meaning of ones and tens - is considered non-straightforward, though it cannot be considered challenging.

3. **Learning gains seen across classes is slightly incremental and not a large jump.** Student performance in common questions that were used to check learning gains across classes showed that performance increased as students move from class 4 to 6 to 8 in both language and maths. However, the extent of improvement was often slightly incremental and not a large jump as one would expect. In most cases, nearly 40% of students in class 8 do not seem to have acquired class 4 competencies. Many misconceptions that students have in lower classes still continue in higher classes and in some cases were found to even become stronger.
4. **Students find it difficult to express their thoughts in their own words in writing.** Their writing does not go beyond the most basic, tried and tested formulaic sentences they probably trained for while in their lower classes. Their writing shows that they are learning language more as a subject and less as a means of natural communication. The writing has a number of errors in spelling and grammar, and punctuation marks are conspicuous by their absence.
5. **In all the states tested, fewer students were found to comprehend what they read.** For e.g., in the oral reading test, more than 87.1% of students in Gujarat could read a simple word, but only 40.3% could read a short passage and only 22.6% could comprehend the information implicit in the passage they read.
6. **There are significant state-wise differences in student performance.** Based on the relative performance/rank of the states in different classes and subjects, an attempt has been made to consolidate the performance of 17 states. Kerala, Maharashtra, Orissa and Karnataka are clearly performing overall better than the national average. Jammu and Kashmir, Madhya Pradesh and Rajasthan were among the states that ranked among the bottom three overall. Bihar performed the same as national average.

A high correlation of 0.85 and above was observed in the performance across the classes and subjects tested in each state. This means that if a state does well in a class or subject compared to other states then it more or less tends to do well compared to other states in other classes and subjects too.

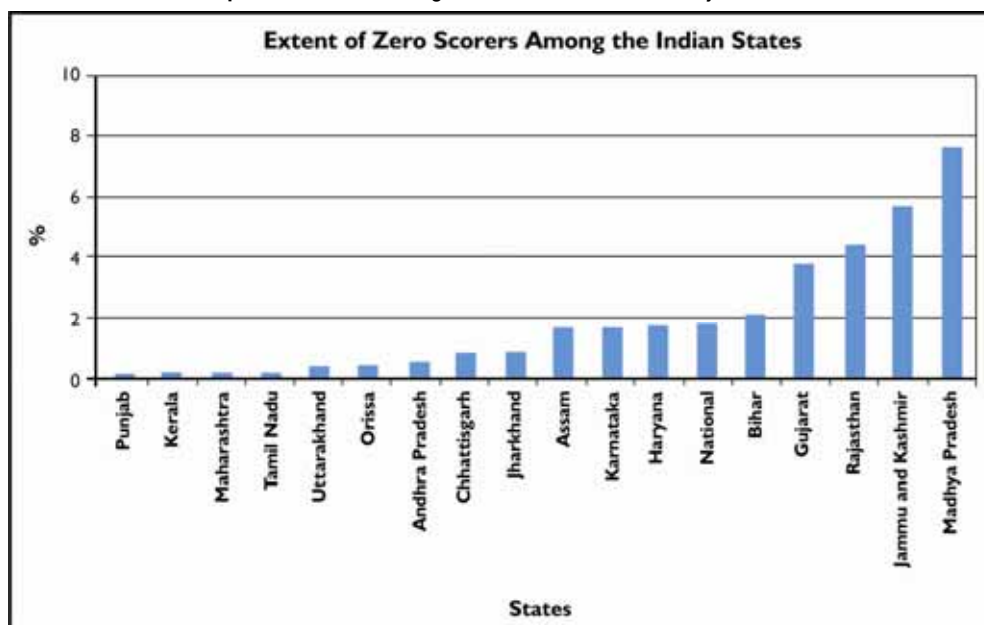
Graph 1: Composite performance index of schools of different states



The union territory of Chandigarh, although tested, has not been included in this comparison. Delhi, though a state, is also not included in this comparison as only class 4 was tested here.

7. **Extent of students scoring zero and the overall performance of a state showed differences.** Andhra Pradesh, Assam, Chhattisgarh and Jharkhand, although below the national average, have fewer students not scoring in the test compared to better performing states such as Haryana and Karnataka, indicating that while the states' efforts could be addressing the lowest ability students, it does not provide adequate support for overall improvement of all students. Jammu and Kashmir, Madhya Pradesh and Rajasthan which ranked among the bottom 3 in overall performance also had the highest number of students scoring zero indicating that these students are being left behind in these states.

Graph 2: Students scoring zero in the classes and subjects tested



The union territory of Chandigarh, although tested, has not been included in this comparison. Delhi, though a state, is also not included in this comparison as only class 4 was tested here.

Can different studies report different performance rankings for states?

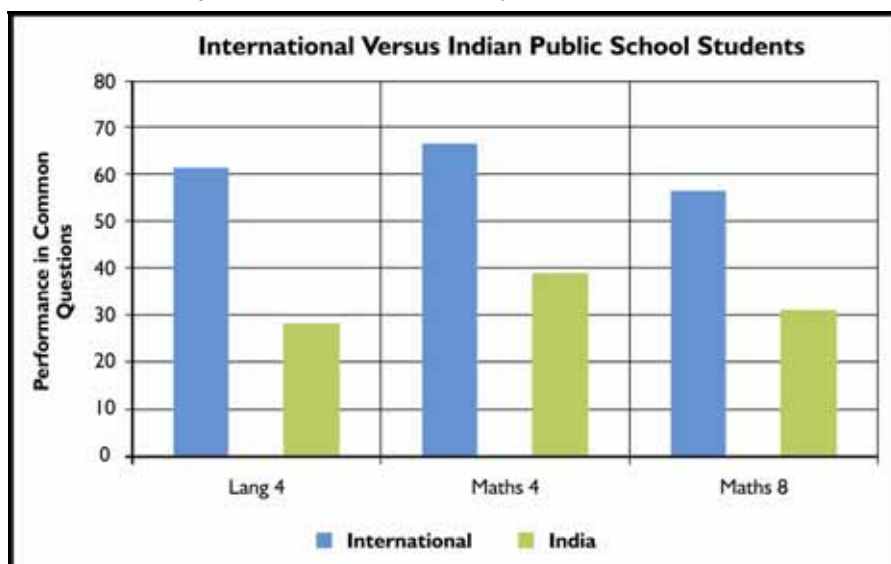
Measuring student learning is a complex exercise and performance reported is influenced by a number of factors. Assuming that the student samples are selected by robust procedures in all studies, test scores would further depend on -

- What is measured?** Test papers used in different studies may not measure identical things. For example, questions could measure rote learning or questions could measure deeper understanding of concepts; and there may be differences in the length of the paper, for a longer paper may mean that more information has been obtained to arrive at conclusions.
- How a test is administered?** The way the test is administered is very important to come to conclusions about student learning. For example, different tests may be administered in a classroom or outside a school environment, testing may be carried out by the teacher who is close to the class or evaluators who have been specifically trained for standardised administration, teams in different states may be synchronised to follow identical processes or different state teams may follow processes that have inherent differences while administering the tests, etc all of which will bring about changes in the test scores.

8. The levels of learning of Indian students in government schools in class 4 and class 8 tested is much lower than the international average as represented by studies like Trends in International Maths and Science Study (TIMSS) and Progress in International Reading and Literacy Study (PIRLS). Similarly, on common questions used from an extensive study for private schools² in India catering to the elite and upper middle class, students in government schools showed a much lower performance.

² Questions were taken from ASSET, a diagnostic assessment test by EI, in which more than 4 lakh students participate from all states of India

Graph 3: Performance in common questions from International Tests



Class 4 Language

Passage Excerpt:

When Lakhan discovered that he had mice in his house, it did not bother him much at first. But the mice multiplied. They began to bother him. They kept on multiplying and finally there came a time when even he could stand it no

Why did Lakhan want to get rid of the mice?

- A. He had always hated mice.
- B. There were too many of them.
- C. They laughed too loudly.
- D. They ate all his cheese.

Question involves retrieving stated information in the text and making straight forward inferences from it.

Sample Question3: Internationally, **79.0%** of students answered this correctly, while **41.8%** of students from Indian public schools could get this correct.

Is it fair to compare school systems that are known to have wide differences in resources and types of students it caters to?

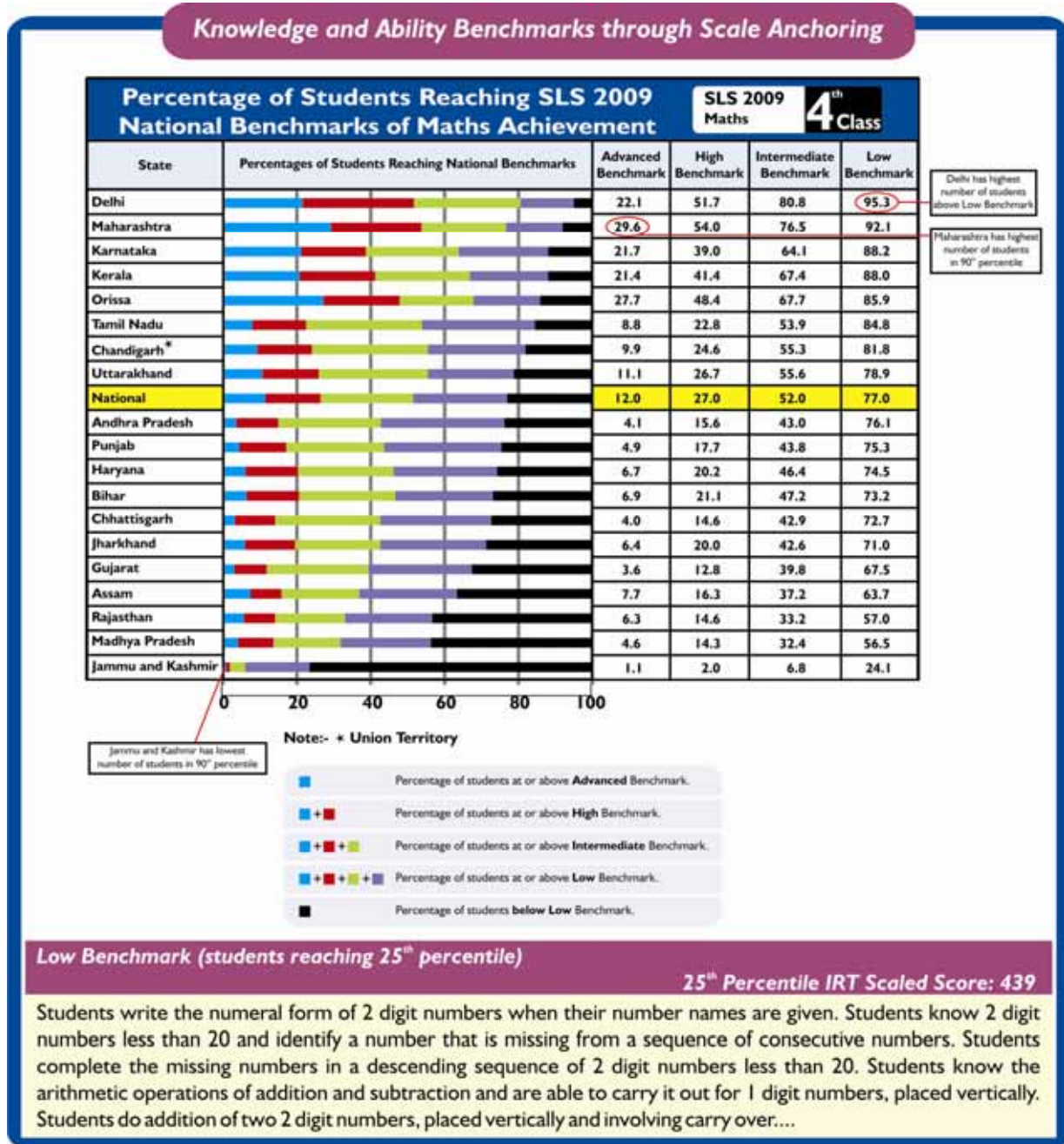
It is well known that schools in other countries as well as the Indian private elite schools may have richer resources with students from better socio-economic strata, which by themselves may contribute to differences in student learning. When such differences exist, it may not be prudent to compare the outcomes from such systems. Hence the purpose of these comparisons should not be to measure or rank the school systems in terms of performance. However, given the goal for our government school system to achieve the best quality education for students, such comparisons do provide the aspiration level and an idea of the performance level students in these schools could aspire for.

9. The comparative performance of boys and girls is similar to many international studies – boys seem to do better in maths, although the SLS study reveals that it is a meaningful difference that matters in class 8 only.
10. Analysis of Background factors showed that when students perceive themselves as being good at studies and think of school as a place of fun and learning, they show good performance. The analysis also reveals the importance of inculcating reading habits in student, for spending 30 minutes each day to read material other than textbooks brings about higher achievement. Teacher training is important especially for mathematics teachers for better achievement of students in the subject. Unexpectedly, analysis shows that students whose teachers have an academic degree developed lower scores in both language and mathematics than those who did not have these.

WHAT IS SO SPECIAL ABOUT SCALE ANCHORING?

Classroom experience - as well as assessment results - clearly show that students perform at different levels of achievement. Students can even be grouped based on their levels of performance. Neither of these, however, provides us concrete steps on what can be *done* to help students performing at lower levels. If we could know firstly, *whether* there are specific topics or concepts which students at lower levels of learning are systematically answering incorrectly, and if so, *what* they are, then remediation could focus on those topics.

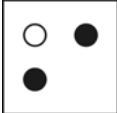
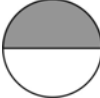
Graph 4: Performance of different states in Class 4 Maths Benchmarks



That is what Scale Anchoring provides, in an extremely scientific manner. The results clearly show which concepts or topics are understood only by students performing at higher levels and which topics are understood by other students also. Certain topics are said to 'anchor' at certain percentile levels of performance (25, 50, 75, 90 percentiles). Insights provided by the Scale Anchoring process can help us understand the way children learn and plan scientific remediation.

Benchmark information can be used in a variety of ways - 1. At the policy level, benchmarks help determine the areas of focus for teacher training (for example, see Graph 4, Tamil Nadu although above national average, has 8.8% of students reaching advanced benchmark in class 4 maths showing the need for training to enable teachers handle concepts anchoring at advanced benchmarks); 2. At the level of curriculum and pedagogy, benchmarks determine patterns in learning that are useful in refining textbooks and teaching learning materials (for example, see Table 1 which shows patterns on how the same topic is acquired in different ability levels); 3. At the school level, benchmarks enable the teacher to bring about classroom remediation. (for example, based on Table 1, a teacher gets to know what are the gaps she should address for her students to reach advanced benchmark in the topic)

Table 1: Exploring Acquisition of a Concept at Different Levels of Student Ability

Class 4 Maths - Concept of a Fraction		
Low Benchmark (25 th Percentile)	If a watermelon weighs 10 kg, how much will half the watermelon weigh? _____	Students understand fractional quantities such as half written in a word form as one out of 2 parts and apply them practically in their daily context.
Intermediate Benchmark (50 th Percentile)	In which figure are one-half of the dots black? Tick (✓) the answer.  A.  B.  C.  D.	Students understand the concept of half as a number divided by 2, for example, in a group of same objects, they know that 3 out of 6 is half the number of that object.
High Benchmark (75 th Percentile)	Which figure is divided into four EQUAL parts? Tick (✓) the answer.  A.  B.  C.  D.	Students understand parts of a whole and can visually identify equal parts.
Advanced Benchmark (90 th Percentile)	Which figure is $\frac{1}{2}$ shaded? Tick (✓) the answer.  A.  B.  C.  D.	Students understand half represented as a fraction and understand it as one out of 2 equal parts and are able to identify the correctly shaded figure based on this.

INSIGHTS FOR CURRICULUM AND PEDAGOGY THROUGH SCALE ANCHORING BENCHMARKS

RECOMMENDATIONS:

- Orienting policies to focus on learning outcomes:** While the Central Government initiated Sarva Shiksha Abhiyan (SSA) in 2002 and set the targets of universal primary education in 2007 and Universal Elementary Education (UEE) by 2010 respectively, it has not set itself a target based on quality of student learning outcomes. It is important that all policies are built with a focus on the learning outcomes, for, if students are not learning, then this defeats the purpose even if all other goals are met.
- Making Low Stakes Diagnostic Assessments – a regular feature of the state programs:** Assessments tend to work well when they are low-stake. Their purpose is to inform students, teachers, schools and even the larger society where schools stand. Without official pressure, the purpose of the assessment is two-fold: to provide support and information, and this itself leads to the creation of a positive peer and/or self pressure. The tests must be created by a body or organisation that is truly independent and is not in a way responsible for providing quality education. Low stakes assessments thus developed should be diagnostic and check for understanding and not simply rote or recall.
- Extensively use Benchmark data from the study to build reform:** A powerful goal of these assessments is to obtain a clear picture on where students and teachers stand with respect to peers in the state, peers in the rest of the country and peers internationally. These are not just numbers, but detailed statements of strengths and weaknesses which can lead to specific action points.

What are Learning Outcomes?

Learning outcomes are statements of what a learner is expected to know, understand and be able to demonstrate after completion of learning. They are essentially student centred or learner centred. They seek to describe the student's learning progress in terms of the knowledge acquired, the comprehension of that knowledge, the capacity to apply it, the capacity to analyse, synthesise and evaluate. Learning outcomes guide the selection and coordination of appropriate content, learning activities, and assessment strategies that promote the overall learning process. Quality of student learning can be monitored against the expected performance for these learning outcomes.

An example of a Learning Outcome in Reading: Students' listen to or read the various types of texts for information, comprehension, and literary appreciation.

In Classes K-4, students should know and be able to do the following:

- Listening to /and reading a variety of Indian and non-Indian literary (class-appropriate fiction, folktales, fables, funny stories, rhymes, plays, diaries, etc.) and non-literary text (simple informational text, text books of other subjects, picture books, simple descriptive/ narrative text, children's encyclopaedia, instructions from science experiments, factual recounts (news stories), lists, etc. - From 'Learning Standards' by Educational Initiatives

How do you check for the quality of attainment in this learning outcome?

These are described in the 4 performance levels – Basic, Partially Proficient, Proficient, Advanced.

e.g. Proficient Level

At this level, students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts. Some examples for 'Proficient Level' performance are:

- Explain or recognize how the author's purpose affects the interpretation of a reading selection.
- Analyze and describe the characteristics of various types of literature

- 4. Student Progress Tracking System:** We have reached a stage in our development where accuracy of available data alone can make a significant difference to our development. This brings to the fore the need for computer-based systems like Student Progress Tracking System in helping to achieving universal primary education nationally. Student Progress Tracking System is a computerised system with unique student identification and aids in tracking student learning / progress on both scholastic and non scholastic domains. Such a system will allow drilling down of information to the level of the individual student and teacher and enable targeting the remediation.
- 5. Large Scale Awareness campaign to redefine attitudes towards learning - a movement against rote learning and for Learning with Understanding:** Rote learning can deceptively look like learning and be mistaken for it. A consensus needs to be gradually built that *rote learning is not learning at all*.
- 6. Widely disseminating the findings of this report among teachers and others:** We recommend that each state make a systemic and detailed plan to disseminate this report, giving every teacher an opportunity to see it, understand and even question and discuss it. The purpose of all this is two-fold: 1. initiate the rote versus learning with understanding debate in the country; and 2. get teachers to start thinking about 'what are children learning and not learning' rather than just focussing on 'what are we teaching or what to teach'.
- 7. Providing effective teacher support based on the feedback from the assessment:** Targeted capacity building in teachers is often enabled by detailed diagnostic assessments such as the SLS, which reveal absolute performance and trends showing weaknesses in groups of students, schools and sometimes even a region or in the entire state. The background factor analysis in the study reveals that teachers with an academic degree were ineffective in achieving higher scores for their students in language and maths while teachers with teacher training qualification were effective in achieving higher scores for their students in maths. This is an indication for reviewing the teacher recruitment and training system.

The future steps would involve dissemination of the findings at the national and state level to enable policy and educational reform.

Chapter 1. DESIGN OF THE STUDY



1.1 Need for the Study

World over, governments and citizens alike seem to agree that good quality education (what and how well students learn) rather than mere school attainment – is powerfully related to improving their individual outcomes in the labour market, enhancing economic growth of a nation and reducing inequality in society (Hanushek and Woessmann 2007; Vegas and Petrow, 2008; World Bank, 2007). While education in the past meant largely the 3R's, the needs of the knowledge society call for critical thinking, learning how to learn and problem solving skills to be imparted to students.

Improving an education system to achieve this quality in education almost certainly requires a multi-pronged effort. Parallel to these efforts, however, it is necessary to establish some method to objectively measure the level of student learning! Further, this method of measuring student learning must be *valid, periodic, external and largely accepted*.

Since the 1960s, some countries like Australia, Belgium, England, Finland, France, Germany (FRG), Israel, Japan, Netherlands, Scotland, Sweden, and United States³ have assessed their students' academic performance regularly through external benchmarking tests to provide a basis for further action.

The need for large scale benchmarking studies that provide granular information across the education system and insights into some very fundamental questions such as - '*Children are going to schools and are being taught*' - *but are they really learning and understanding what is taught to them? How much are they retaining? Is learning happening at a superficial level? What do children at different levels of the attainment know and are able to do? Are children of some states and districts doing better than others, if so, how?*' assume greater importance in our collective striving to provide quality education to our children.

Educational Initiatives has itself conducted benchmarking studies in India. Two studies conducted within the last 3 years include the Student Learning in Metros Study (2006) and a benchmarking of student learning in local body schools in 30 towns of the country (2007). This is the third large scale study and covers urban and rural schools in 18 states of India.

1.2 About the Study

'Student Learning Study (SLS)' is a benchmarking study of student learning conducted by Educational Initiatives (EI), with the support of Google org, USA. The study has been carried out in 48 districts in 18 states and 1 Union territory of India. About 101643 students from 2399 selected government schools studying in Classes 4, 6 and 8 were tested in. language and maths. The test papers for each class were developed commonly in 13 languages.

The goals of the 'Student Learning Study' are as follows:

- To provide independent 3rd party benchmarking of student learning for use by policy makers and researchers both at the Central and the State levels.
- To provide insights into comparative performances of different states and to create cross-learning and remedial opportunities.
- To establish student learning in terms of knowledge and skills and feedback on the learning gaps, common errors and misconceptions.
- To provide pointers for further research.
- To diagnose the learning issues at the systemic level.

In short, Student Learning study is a **sophisticated and comprehensive** study that aims to provide detailed question-wise and state-wise data for monitoring learning across classes, and provide detailed diagnostic feedback. However, at this point, it has to be recognised that well designed and well implemented benchmarking achievement studies by themselves may not bring about the necessary systemic shift to improve education, but they can provide powerful 'data-driven' insights into the existing learning gaps for students to learn better, teachers to teach better, and schools and policy makers to operate more effectively.

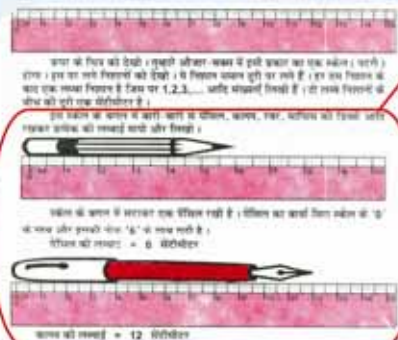
³ The First International Mathematics Study (FIMS); <http://www.iea-hq.org/fims.html>

Textbook Case Study - I



The state text books of class 3 and 4 shows how the length of an object can be measured with a scale. However, we see that a lot of students are having trouble measuring the length of an object correctly even at classes 4, 6 and 8. Some of the excerpts from different state textbooks are shown below:

What the class 3 textbook says...



Measure and write the length of a pencil, pen, rubber/eraser, matchbox etc. by placing them beside the scale.

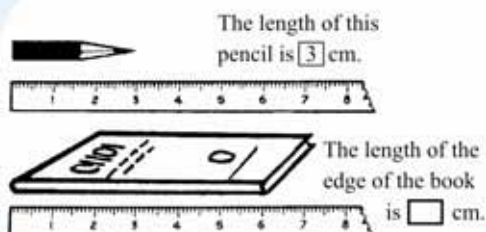


There is a pencil next to the scale. The left part/side of the pencil is kept on the '0' of the scale and its tip is next to '6'. Length of the pencil = 6 centimeters



Length of the pen = 12 centimeters

Source: Bihar, Chapter – 20, Page – 111, Class - 3



Source: Maharashtra, Class 3, Section 3, Page 59

वर्षा ऋतु पट्टिका का मुख्य अंक मिर चुका हो या माप-माप बिगड़ नहीं दे रहा हो भी हो इस रेखा-वर्णन की सहायता प्राप्त कर सकते हैं। नीचे दी गई बाइनिंग में एक स्थान में रेखा-वर्णन ज व को बिलाने के लिए पट्टिका इस प्रकार रखी है।



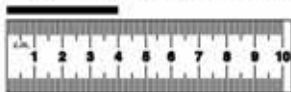
क्या आप रेखा-वर्णन ज व की सहायता प्राप्त कर सकते हैं ? हाँ, इस रेखा-वर्णन ज व की सहायता प्राप्त कर सकते हैं। कैसे ? व देखते हैं कि पट्टिका पर जो बिन्दु मारे ज तथा व को जुड़े हैं बिन्दु 3 और 10। इसलिए अन्तर 10-3 = 7 रेखा-वर्णन ज व की सहायता की दृष्टिगत है। अतः रेखा-वर्णन ज व 7 सेंटीमीटर लम्बी है।

Source: Punjab, Class 3, Chapter 8, Page 130

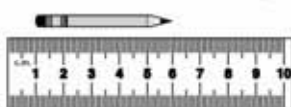
We asked Class 4, 6 and 8 students...

Class 5 English

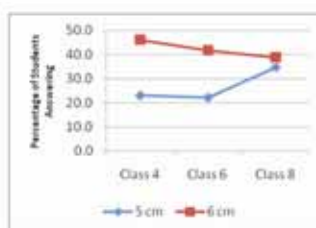
The length of the line in the figure below is 4 cm.



How long is the pencil shown in the picture? (Use the ruler shown in the picture.)



cm.



Class	5 cm	6 cm
4	23.0%	46.0%
6	22.1%	41.7%
8	34.7%	38.8%

It looks like a majority of children (46.0% in class 4, 41.7% in class 6, 38.8% in class 8) are learning measurement mechanically and wrote 6 cm as answer. When measurement does not start from zero, as is in this question, they do not see a need to subtract the starting point from the final one (in our case, $6 - 1 = 5$ cm).

1.3 Salient Features of the Study

The study is the biggest of its kind in any country of the world and has a number of firsts to its credit:

Independent Benchmarking: This is the first time in India that an independent, benchmarking assessment is entirely carried out by a single external agency which has detailed benchmarking information from past studies.

Partners: Expert partners in this study include professors from - the Indian Statistical Institute (*on sampling*), NUEPA (*on providing school lists with enrolment*), CIIL (*for test adaptations and translation validations*), HBCSE (*for feedback on test items*), University of Michigan (*for expertise on collection of background parameters and checking their impact on learning*) and other international experts for inputs based on NAEP and TIMSS sampling, IRT scales for embedding international anchor items and validation of IRT data interpretation on specific items.

Coverage: About 150000 students were sampled from 2399 government run schools in 48 districts in 18 states and 1 Union Territory. The coverage included samples from approximately 74% of the Indian population from urban as well as rural India. The actual number of test takers was 101643, representing a presence rate of 66% which is more or less in line with the regular presence rate found in government schools.

Scientifically Designed Testing Cycle: The study followed a detailed scientific process of test development supported by curriculum checking; pilot tests in 3 states to establish validity of items; statistical and qualitative analysis of pilot tests; further fine tuning of the final papers; detailed test adaptation, translation validation and harmonisation of papers across languages; master trainer workshops, training of evaluators; standardised test administration; data entry with evaluator coded OMR sheets; and diagnostic data analysis prior to releasing reports in the public domain.

Specially Assembled Question Paper: Test items for classes 4 and 6 were drawn extensively from the earlier Municipal School Benchmarking study in which the items were tested on 35000 students across 30 urban towns of India. Some Class 4 and Class 6 items and all of Class 8 items were developed anew for the study. In a move to benchmark achievement of Indian students with those from other countries, some items from international studies such as the Trends in International Mathematics and Science Study (TIMSS) and Progress in Reading Literacy Study (PIRLS) were used as anchor items for comparison. TIMSS class 4 and 8 items and PIRLS class 4 items were used for this purpose.

Language Oral Reading Test: The Language oral reading test was done with a fewer sub sample of students on a one-on-one basis at class 4 level in all the schools tested. The test had a total of 8 questions and required students to read aloud words and sentences of different difficulty levels. The students also had to read a short passage and then answer questions based on information stated explicitly and implicitly in the passage. This was done to gain insights into whether the students could decode and comprehend basic information from what they read.

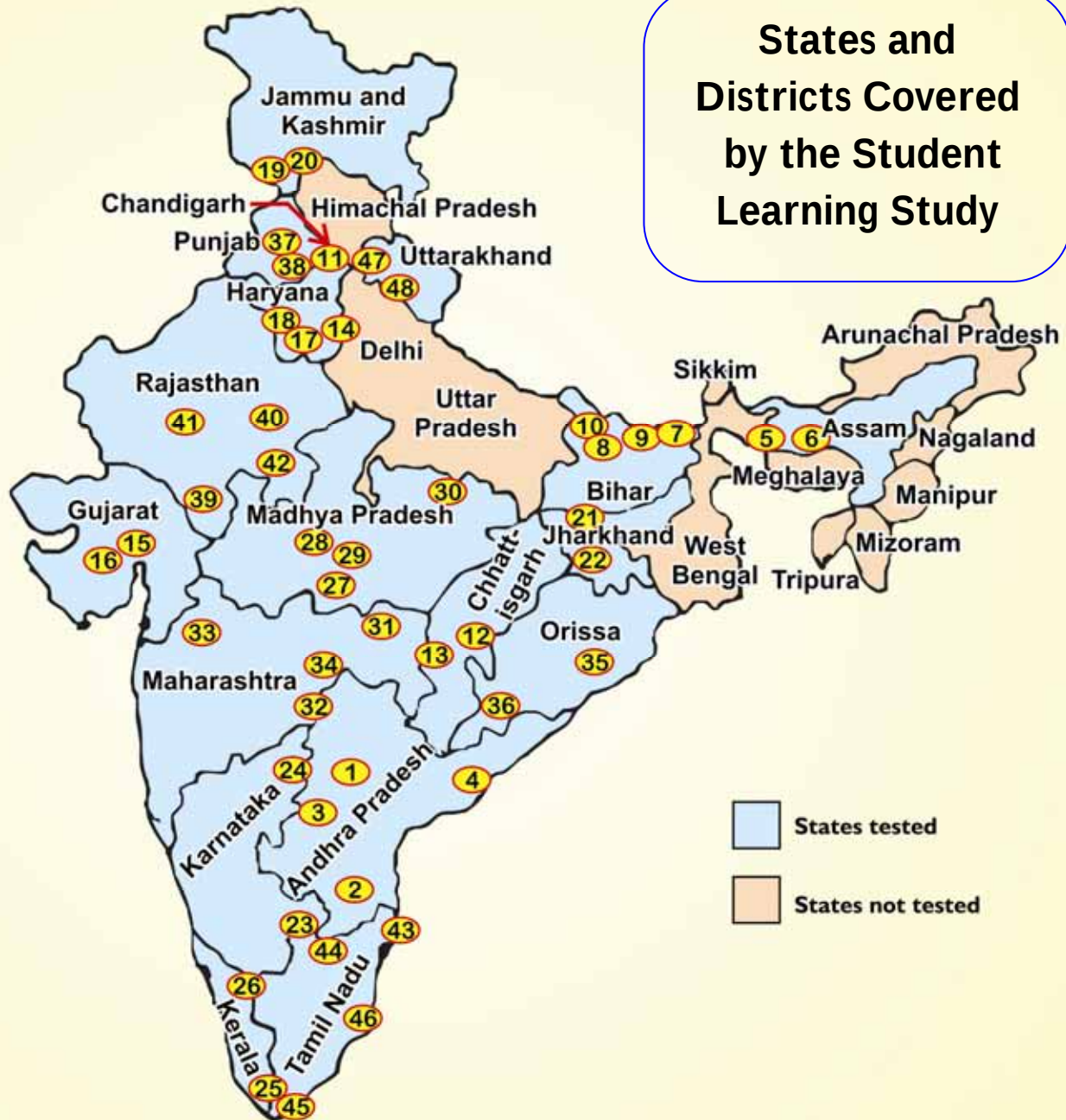
Multiple Languages: The tests covered 13 languages. The test adaptations and translations followed a rigorous process for compliance to the international test commission guidelines (ITC). Comparability of the different language versions were established through harmonisation of all papers in an iterative process to account for cultural and linguistic adaptations, reverse translation, multiple validity checks by a team of language experts, documentation of the adaptations, etc.

Specially Trained External Test Administrators: The test administration was standardised completely by use of a special team of evaluators who were trained rigorously to cover aspects of test administration that included reading out the 'group oral' question items, pacing the test as well as assigning appropriate codes for the answer responses given by students.

Benchmarks through Scale Anchoring: For the first time in India, **Knowledge and Ability Benchmarks** on what students know and do at different ability levels is made available through the advanced scale anchoring process. The scale anchoring process used the modern item response theory (IRT) as per international practice to arrive at item parameters and test parameters.

Analysis: Different types of analyses were carried out on the collected data to extract patterns in performances and to understand differences in learning levels across different groups. Advanced statistical methods were used to confirm patterns of learning. Distracter analysis enabled identification of misconceptions and common errors. A small sub sample of language answer scripts in Hindi were checked by a group of experts to gain insights into the ways in which students perform on the free response questions that required students to write a few sentences.

States and Districts Covered by the Student Learning Study



LEGEND

Andhra Pradesh

- 1 Hyderabad
- 2 Kadapa
- 3 Mahabubnagar
- 4 West Godavari

Assam

- 5 Dhubri
- 6 Kamrup

Bihar

- 7 Araria
- 8 Patna
- 9 Saharsa
- 10 Saran

Chandigarh

- 11 Chandigarh

Chhattisgarh

- 12 Raipur
- 13 Rajnandgaon

Delhi

- 14 Delhi

Gujarat

- 15 Ahmedabad
- 16 Surendranagar

Haryana

- 17 Gurgaon
- 18 Rewari

Jammu and Kashmir

- 19 Jammu
- 20 Kathua

Jharkhand

- 21 Chatra
- 22 Ranchi

Karnataka

- 23 Bangalore
- 24 Gulbarga

Kerala

- 25 Thiruvananthapuram
- 26 Waynad

Madhya Pradesh

- 27 Betul
- 28 Bhopal
- 29 Hoshangabad
- 30 Panna

Maharashtra

- 31 Nagpur
- 32 Nanded
- 33 Nashik
- 34 Yavatmal

Orissa

- 35 Khordha
- 36 Koraput

Punjab

- 37 Ludhiana
- 38 Sangrur

Rajasthan

- 39 Dungarpur
- 40 Jaipur
- 41 Jodhpur
- 42 Tonk

Tamil Nadu

- 43 Chennai
- 44 Dharmapuri
- 45 Kanniyakumari
- 46 Pudokottai

Uttarakhand

- 47 Dehradun
- 48 Tehri Garhwal

Background Questionnaires: Background information about the student's school and home environments was collected with the help of school, head teacher, teacher and student questionnaires to identify factors associated with learning. The data collected can be broadly grouped into questions on school background, teacher background, teacher perception, head teacher background, student background and student perception. The student questionnaire was administered in the classroom after the main tests. The school, head teacher and teacher information were collected by the district coordinators through one-to-one interviews in the school.

Diagnostic feedback: The study provides detailed diagnostic feedback at item and paper level to schools and states and suggests remedial actions to the different stakeholders.

Video Study: Following the assessment, student interviews were conducted in 3 states – Rajasthan, Gujarat and Andhra Pradesh to understand why students answer certain questions in the ways they do. The student interviews covered 2 concepts and common errors each in maths and language from classes 4 and 8. These interviews were video-recorded and provide a powerful window for stakeholders to understand how and why children think in certain ways.

1.4 Study Coverage

The study initially envisaged covering schools across rural and urban areas in 21 states of India. However, the study finally covered 18 states and 1 union territory of India based on the permissions that were received. Uttar Pradesh, West Bengal and Himachal Pradesh did not agree to participate in the study. In Delhi, classes 6 and 8 refused to participate while class 4 participated. Chandigarh, although a union territory, was also the capital of the states of Punjab and Haryana and hence was included in the study. Jammu and Kashmir government wanted the study to cover the areas of Kashmir too, however due to regional tensions during testing the study covered only districts in Jammu.

48 districts were selected on the basis of sampling design (described in chapter 6 of the report) as representative of the states to take the tests. In all the states, schools in both urban and rural areas were covered. The number of schools and students who took part in the tests is as below:

Table 1.1

Class	Number of Schools Tested		Number of Students Tested	
	Language	Maths	Language	Maths
4	1411	1410	29578	29513
6	1383	1381	35945	35604
8	1300	1301	36120	35967

1.5 Question Paper Design

The heart of this study is definitely the instrument design. A review of most similar tests conducted in India suggests that tool design - specifically design of the achievement tests - has not been given the importance it deserves. Many important studies, even when published into books, often do not enclose or append the actual test, though the subsequent data analysis is highlighted a lot.

The challenges in the development of tools in this study were - the need for the papers to be valid across 18 states and 13 different languages; the low levels of learning known to exist among these children and the need for the tools to accurately measure learning at these levels. Our prior experience in achievement testing in government schools of India also showed that students, especially in classes 1 to 5, often find it difficult to read questions on their own and this limits their capability to answer the questions. Hence the structure of the tests was designed in such a way that the measurement of other abilities is not impacted by the reading difficulty faced by students.

Paper Format: The test papers included three different components - *Group Oral, Written and Individual Oral*. The 'Written' test had questions that were read and answered by students themselves. The 'Group oral' part of the test had questions that were read out orally by the evaluator and the students were asked to respond by writing the answer in the paper. The individual oral test was a language reading test done with a fewer sub-sample of students on a one-on-one basis in each class. Here the evaluator administered the item orally to each student who in turn responded orally. The evaluator captured the oral responses given by student in the individual oral test in the form of a set of predefined answer codes.

Distribution of Questions in Each Paper and the Duration:

Table 1.2

Class	Papers	Number of Questions			Duration (mins.)
		Written	Group oral	Total	
4	Language Test	24	13	37	120
4	Maths Test	32	10	42	120
6	Language Test	35	8	43	120
6	Maths Test	47	0	47	120
8	Language Test	48	0	48	120
8	Maths Test	52	0	52	120
4	Language Oral Reading test	Completely Oral		8	10 minutes per student

Competencies Tested: The student learning study sought to assess students' understanding and ability to apply what they had learnt. The questions were not based simply on the ability to recall information or use formulae or procedures, as most school exams are, but also tried to check if the students have understood and internalised the concepts. Each question was assigned a specific competency and each competency had a minimum of not less than 3⁴ questions. The National Curriculum Framework documents, the MLLs, the State textbooks, as well as Standard international frameworks of tests like the TIMSS and PIRLS and EI's large scale benchmarking studies for the government schools were referred to while finalising the competencies for each paper. Class 4 and Class 6 tests was built largely on the previous large scale municipal schools benchmarking study of EI covering 30 Urban towns across 5 states of India. The list of competencies tested in each paper is listed in Appendix A. A description of these is outlined below:

I. Language: Language is important as a means of communication, learning, advanced expression and appreciation. It also defines one's identity and is a vehicle of culture. The test papers assessed 7 competencies in class 4, 8 competencies in class 6, and 6 competencies in class 8. These focussed on vocabulary, appropriate usage, aspects of reading comprehension, spellings, and basic aspects of writing such as reorganizing a jumbled sentence or sentences, writing a single to a few meaningful sentences on a picture or a topic, punctuating a sentence, completing a cloze passage where words had to be filled, completing a dialogue between a shopkeeper and a girl who goes there to buy pencils, and completing a miniature short story by adding 2 sentences from the word clues given. The passages used were unseen and included short stories, descriptive text, authentic material such as bottle labels, school notices for a competition, the table of contents of a novel, etc. In Class 4, a short story with 4 reading comprehension items from the international 'Progress in Reading Literacy Study (PIRLS)' was included. A slightly longer short story released by PIRLS for class 4 was included in the class 6 and 8 papers to check if higher class students in India are able to do these.

II. Mathematics: Maths should help develop the child's resources to think and reason mathematically, to pursue assumptions to their logical conclusion and to handle abstraction. The key areas in primary and middle school Maths - numbers and operations, fractions, decimals, ratios, percentages, measurement and estimation, data interpretation, algebra, geometry and problem solving were covered. A few items from the international 'Trends in International Maths and Science Study (TIMSS)' for class 4 and 8 were included in the papers to provide a relative understanding of the performance we find in our schools in India.

Question Types: The questions in all papers included both *multiple choice items* and *free response items*. From an objectivity, analysis and scoring point of view, multiple-choice questions offer many advantages. Our experience also shows that urban students are generally familiar with multiple choice formats made popular in India through various game shows in TV, etc. However, multiple choice items do not offer the scope to capture how the students respond independently when asked to write a few sentences, complete a cloze passage or dialogue, or give a numerical response to a computation, etc. Rural students especially at lower class levels also generally respond better to free response items due to familiarity with the traditional format of questions used in the textbooks. Free response items were scored objectively using a well-defined scoring rubric and well trained evaluators. The tests were competency-based in the sense that every question was linked to a particular competency in the skill framework/test blueprint. Apart from the *competency balance*, the tests were also designed to include '*straightforward*' questions and '*non-straightforward*' questions. A straightforward question is one which

⁴ The only exception was in Class 6 and 8 Language - in the skill "Expresses ideas coherently on a given topic by writing 4 sentences and completes description of an incident/story using given word clues" which had only 2 questions each.

has a 'form' as it appears in the textbook. A 'non-straightforward' question is not very different or challenging⁵, but provides information on whether children are able to perform only textbook type problems or whether they are able to apply the essential skills and competencies and use their understanding to write answers.

Types of Questions used in the Tests: The types of questions used in the tests are illustrated below:

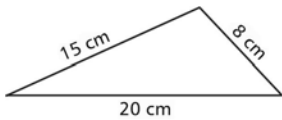
A. Questions check for deeper understanding of concepts

Class 6 Language Which words below best describe the wolf? A. greedy and clever B. innocent and hungry C. foolish and greedy D. clever and hungry <i>Sample Question 1: This question checks whether students have understood the character described in the story and are able to evaluate the different descriptions to choose the one that fits the character.</i>	
Class 4 Maths Fill in the appropriate number in the box. $3 \times \square = 3 + 3 + 3 + 3$ <i>Sample Question 2: This question checks whether students have really understood the concept of multiplication. The question tries to find out if children are able to find patterns between the algorithm and its meaning.</i>	Class 6 Maths In which of these numbers below does 3 have the greatest place value? A. 136 B. 239 C. 301 D. 743 <i>Sample Question 3: This question checks whether students have understood the concept of place-value. The options are constructed in a manner to capture what a student understands while reading such questions. For e.g. a student choosing option D would seem not able to differentiate between the terms - greatest place value of a particular digit and the greatest number.</i>

B. Questions check for learning that is straightforward or text-bookish

It is sometimes argued that while students may not be learning with proper understanding, they can do extremely well on question types they are familiar with. In order to check this, the test also contained some questions which can be best described as 'typical', 'text-bookish' or 'straightforward'. Here are two examples, one each from maths Class 4 and maths Class 6.

Sample Question 4 and 5 below try to find how well are children able to perform on questions based on textbook. The student response on these questions might give us the information if a child is able to perform the operations based on the rules taught in class.

Class 4 Maths Write the answer. $\begin{array}{r} 39 \\ - 6 \\ \hline \\ \hline \end{array}$ <i>Sample Question 4: This question checks if students are able to do subtraction of one digit number from a 2 digit number without borrow.</i>	Class 6 Maths What is the perimeter of this shape?  _____ cm. <i>Sample Question 5: This question checks if student understand what perimeter is and are able to calculate that from the given values.</i>
---	---

⁵ For example, a question that asks what 4 tens and 3 ones is considered straightforward; whereas one that asks what is 5 ones and 4 tens – in which the order is reversed and which tests whether the child has understood the meaning of ones and tens - is considered non-straightforward, though it cannot be considered challenging.

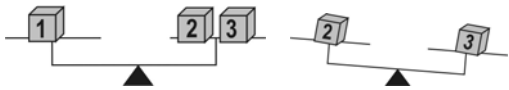
C. Questions check for application of concepts

The application-based questions are basically the ones in which a child is expected to apply different concepts to solve problems based on what is taught in class.

Class 6	Maths										
Look at this calendar. What day will 1st April be? _____	What do you have to do to each number in Column A to get the number next to it in Column B? <table border="1"><thead><tr><th>Column A</th><th>Column B</th></tr></thead><tbody><tr><td>10</td><td>2</td></tr><tr><td>15</td><td>3</td></tr><tr><td>25</td><td>5</td></tr><tr><td>50</td><td>10</td></tr></tbody></table> A. Add 8 to the number in Column A. B. Subtract 8 from the number in Column A. C. Multiply the number in Column A by 5. D. Divide the number in Column A by 5.	Column A	Column B	10	2	15	3	25	5	50	10
Column A	Column B										
10	2										
15	3										
25	5										
50	10										
Sample Question 6: In this question, the student needs to understand the logic and design of a calendar. The calendar shown is for the month of March, whereas the question is on the following month, requiring the student to extrapolate the information and apply it appropriately.	Sample Question 7: In this question, the student has to look for relationship patterns between the numbers in column A and B. The student needs to apply her understanding of whole number operations in order to deduce the relationship.										

D. Questions check for reasoning ability

The following questions test the reasoning abilities along with basic understanding of other concepts also.

Class 6	Maths
Three boxes: 1, 2 and 3 are weighed as shown below.  The heaviest box is _____. The lightest box is _____.	Do you think the lion liked the hare? What happens in the story that shows this? _____ _____
Sample Question 8: The question tests if the child is able to understand the concept of weights and then reason out based on the information given.	Sample Question 9: In this question, the student has to reason and support his argument with evidence from the story.

E. Passage questions

Most questions in the language papers are based on 'unseen' passages and authentic material seen in daily life such as tables, labels, notices, etc.

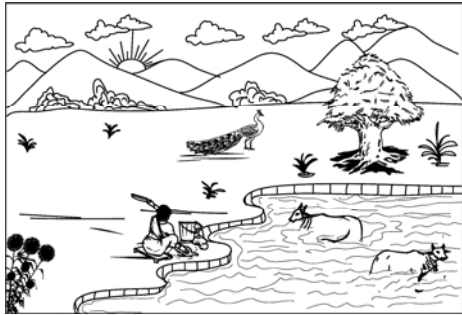
Class 8	Language
What is the main message of this story? A. Run away from trouble. B. Check the facts before panicking. C. Even lions that seem kind cannot be trusted. D. Hares are fast animals.	The pickle should not be eaten after the month of _____ (write the month and the year)
Sample Question 10: The question follows a passage about a Hare who was scared of earthquakes and tests a student's ability to gain an overall understanding of the passage and determine its theme.	Sample Question 11: This question follows a passage displaying a label from a pickle bottle. The label provides information on the ingredients of the pickle, the cost, date of expiry, manufacturer details, etc. The student has to infer that the pickle is not usable beyond the expiry date.

F. Questions check for holistic language learning and basic writing ability

Cloze items are emerging as holistic test items to test language learning. The language papers used both restricted and open cloze items. In addition to general reading and language skills, a cloze test requires the following abilities

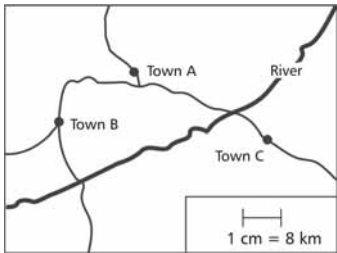
- knowledge of vocabulary
- knowledge of collocation⁶ and colligation⁷
- attention to nearer and farther context
- flexibility as to word and context interchangeability

The test papers also asked students to write a few sentences on a given topic or picture.

Class 6 Language	Class 4 Language
Complete the passage using suitable words of your own. Ganga makes and sells garlands. She buys the flowers from the _____ and she makes them into garlands. She uses flowers of different _____. But she usually makes garlands of small white flowers because they _____ the most. She sells _____ to many people. She sells all the garlands she makes for the day.	Write 3 sentences about this picture. 
Sample Question 12: The question is a cloze item that checks for student's ability to complete the passage appropriately using words of her own.	Sample Question 13: The question checks for student's ability to write 3 sentences that are meaningful and connected to the picture. The sentences are also checked for grammatical correctness but not for spellings.

G. International benchmarking questions

In order to be able to compare the students' performance against international benchmarks of student learning, questions were also taken from TIMSS and PIRLS studies. For instance, based on students' performance on the sample question 14 below, it would be possible to compare the level of basic maths conceptual understanding of the Indian students under study with internationally accepted standards.

Class 8 Maths	Class 4 Language
One centimetre on the map represents 8 kilometres on the land. About how far apart are Town A and Town C on the land?  A. 4 km B. 16 km C. 35 km D. 50 km	How does the story show you what the mice thought was happening? A. by telling you what Labon thought of the mice B. by describing where the mice lived C. by telling you what the mice said to one another D. by describing what the mice were like
Sample Question 14: The question requires the student to estimate visually the distance using the scale mentioned below in the map.	Sample Question 15: The question checks for student's comprehension and requires the student to explain based on the textual elements that indicate how the mice in the story were thinking.

Finally, it should be mentioned that trick questions, as well as questions that sought to deliberately confuse or mislead students were scrupulously avoided. To summarise, the questions aim to be relevant, to relate to the topics taught in school and to test whether the student has truly understood the concept, or has only gained superficial knowledge.

⁶ When words are used together regularly, rules are formed about their use not for grammatical reasons, but because of the association. 'Black and white' appear in that order because of collocation; they are always in that order and to put them the other way around seems wrong. For the same reason we 'make a mistake' when we 'do a test'. The reason for using these verbs with these is that we always do; this is collocation

⁷ Colligation is a type of collocation, but where a lexical item is linked to a grammatical one. Surprising, amazing and astonishing are nearly synonymous. We can say it is astonishing/surprising/ amazing, but we tend to say it is not surprising and not the others- surprising colligates with the negative

1.6 Background Questionnaires

SLS study administered questionnaires to students and teachers in order to collect background information and measure relationships and identify key factors associated with student learning. The data was collected on 76 variables that can be broadly grouped into questions on school background, teacher background, teacher perception, head teacher background, student background and student perception. The student questionnaire was administered in the classroom after the main tests. The school, head teacher and teacher information were collected by the district coordinators through one-to-one interviews in the school. The table in the page 49 gives the list of variables for which information was collected in the questionnaires. The questionnaires are provided in Appendix U and V

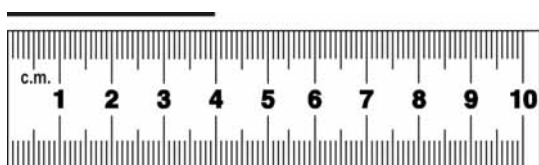
1.7 Video Study

Following the assessment, student interviews were conducted in 36 classrooms across 3 states – Rajasthan, Gujarat and Andhra Pradesh to understand why students answer certain questions in the ways they do. A total of 3 schools were interviewed in each state in the districts of Tonk, Ahmedabad, Hyderabad and Mahabubnagar. In each state, 1 urban school and 2 rural schools were sampled for the student interviews for each question. In each school, 6 questions were interviewed and video recorded. The table below gives the details of the interviews carried out.

Table 1.3

Class	Subject	No of Questions	Number of Schools Per State	No of States	Medium of Instruction
4	Maths	1	3	3	Hindi, Gujarati, Telugu
8	Maths	1	3	3	Hindi, Gujarati, Telugu
4	Language	2	3	3	Hindi, Gujarati, Telugu
8	Language	2	3	3	Hindi, Gujarati, Telugu

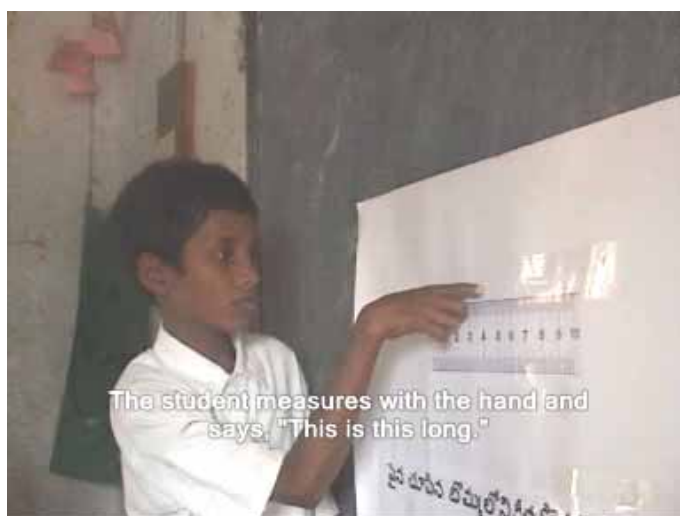
The questions for the interviews were selected on the basis of their performance in the tests, and where the student responses clearly indicated presence of misconceptions. One question each in Maths 4 on measuring a pencil with a scale and in Maths 8 on students understanding of place value of digits in a 3 digit number, were thus selected for the interviews. In language, a passage was selected for reading in class 4 and similarly in class 8. The students were probed further with 2 questions –i) which checked if students are able to retrieve explicitly stated information and ii) which required the students to make connections from ideas mentioned explicitly and implicitly in the passage. The interviews were carried out by trained EI education specialists who probed the students keeping in mind not to prompt or show approval or disapproval while exploring the answers given by the students.



The length of the line in the figure above is 4 cm.
How long is the pencil shown in the picture?
(Use the ruler shown in the picture.)



_____ cm.



A Class 4 student explaining how he understood the question.

Chapter 2. OVERALL FINDINGS



2.1 Main Findings

1. In the lower classes, a fair amount of 'rote-based' or 'procedural' learning is evident in very basic numeracy skills such as number sequencing, operations of whole numbers involving 1 or 2 digits; naming of numbers; reading clock time, understanding currency, etc. Even in procedural learning, students are able to comparatively handle questions that are very 'straightforward' and closer to what one would practise from a typical textbook and not when they are slightly atypical. In higher classes, students are falling behind in all learning, even procedural.

The term 'rote-based' learning is used here in a broad sense to indicate questions that are either typical, straight from the textbook, or is a very familiar question type to the student in its form. Students who rote-learn may be able to handle these questions but not when questions dig a bit more for understanding of the concepts.

Rote-based or Procedural questions can further be distinguished into 'Straightforward' and 'Not-Straightforward' questions. 'Straightforward' questions are those that are closer to what one would see in a typical textbook and hence the student is very familiar with, while 'Not-Straightforward' questions are questions which have been tweaked slightly or have an atypical nature to them.

The results from the different states show that in lower classes, the students are doing relatively well in very basic numeracy skills such as number sequencing, operations of whole numbers involving 1 or 2 digits; naming of numbers; reading clock time, understanding currency, etc if they are asked as straightforward procedural questions.

Class 4 Maths

What time is this clock showing?



Sample Question 16: This is a straightforward procedural question checking if students are able to read clock time. **71.4%** of students answered this correctly.

While students are able to do rote-based questions that are straightforward, they falter when these questions are slightly atypical. The use of zero in the questions below makes these questions unfamiliar and atypical compared to the usual multi-digit subtraction or place value question they may practise in the classroom.

Class 4 Maths

Write the answer.

$$\begin{array}{r} 70 \\ - 43 \\ \hline \\ \hline \end{array}$$

Sample Question 17: This is a 'Not-Straightforward' procedural question due to the presence of zero in the minuend. **49.4%** of students answered this question which involves multi-digit subtraction with regrouping correctly. The number of students getting this question correctly is low considering these are class 4 students and subtraction with regrouping is taught at class 2.

Class 6 Maths

Which of the following is true?

- A. $69 = 6 + 9$
- B. $69 = 6 + 90$
- C. $69 = 60 + 9$
- D. $69 = 60 + 90$

Sample Question 18: This is a 'Not-Straightforward' procedural question testing place value. **52.1%** of students at class 6 answered this question which involves understanding the concept of place value. Place value is taught from much lower classes, and the students also could have used their understanding of addition to get this answer correct. The number of students getting this question correct is low considering these are class 6 students.

Similarly in language, students are decoding the print well to some extent, have basic vocabulary in place and are able to match picture with simple sentences that describe what is in the picture, however, a large proportion of students are not able to use simple interrogative words such as "What", "Where" etc appropriately if the question form is slightly atypical.

Class 4,6 Language

Kheer is _____ to taste.

- A. bitter
- B. sweet
- C. spicy
- D. sour

Sample Question 19: **77.4%** of students answered this correctly in class 4, while **89.6%** of class 6 could answer this correctly.

Class 6 Language

Tick (✓) the picture that best matches the sentence.
A woman is going to the market with her children.



Sample Question 20: **61.7%** of students answered this correctly in class 6.

Class 6 Language

Fill the blank with an appropriate word.

Question: is the book?

Answer: It is on the table.

Sample Question 21: Only **20.9%** of students at class 6 answered this question correctly, which involves understanding the usage of the interrogative that will match the context given. The form in which the question has been framed may be slightly atypical and the number of students getting this question correct is low considering that these are class 6 students.

As students are moving to higher classes, one finds that they are falling behind even in the procedural, straightforward questions that are very familiar to them in a typical textbook. Students seem to have difficulty in almost all the topics such as negative numbers, rational numbers, exponents, averages, ratio & proportion, area, perimeter, discount price, square roots, measurement and algebra.

Class 8 Maths

The ratio of boys to girls in a class is 2:3. If there are 12 girls, what is the number of boys?

Sample Question 22: **16.0%** of students answered this correctly in class 8.

The average mark of 3 children in a subject is 40. If 2 of those children got 41 marks, the 3rd child must have got _____.

Sample Question 23: **15.4%** of students answered this correctly in class 8.

Solve for x.
 $3x = x + 6$

Sample Question 24: **23.1%** of students answered this correctly in class 8.

2. Whatever little learning that is taking place is not ‘Learning with Understanding’ as students find it difficult to answer questions that require a deeper understanding of the concepts.

What is ‘Learning with Understanding’? The answer when one ponders over this is not so obvious. Students can acquire knowledge and recall reams of facts and demonstrate routine skills without understanding their basis or when to use them. However, understanding seems to go beyond knowing something. Knowing and recalling draws largely on memory while understanding seems to go deeper.

A student can be said to have learnt with understanding when she/he is able to apply what is learnt in a different situation in real-life context; is able to solve real-life problems; is able to restate learning in own words; is able to integrate learning from different sources/subjects as needed; and is able to answer questions phrased in a slightly different form.

Why is ‘Learning with Understanding’ important? In today’s world inventions, discoveries, way of carrying on day to day life, gadgets people use, etc are changing at a tremendous pace each decade. Knowledge, as in information is available instantaneously with the click of a button. In such a scenario, what is the use of possessing rote knowledge or routine skills that one does not know where and when to actively use? The key differentiator man brings in is the higher order skills such as critical thinking, creativity, learning to learn, etc. which can develop only when one learns with understanding.

The results of the ‘Student Learning Study’ show that whatever little learning that is taking place is clearly NOT one with understanding. Students were finding it difficult to answer conceptual questions. For example, while 67.1% of students could do a straightforward double digit versus single digit multiplication, only 30.4% of students could actually understand the concept of multiplication as a repeated addition.

Class 4	Maths
Write the answer. $\begin{array}{r} 43 \\ \times 2 \\ \hline \\ \hline \end{array}$	Fill in the appropriate number in the box. $3 \times \square = 3 + 3 + 3 + 3$
Sample Question 25: This is a procedural question that checks for the process of multiplication. 67.1% of students answered this correctly.	Sample Question 26: This is a conceptual question that checks whether the student is able to link multiplication with repeated addition of a number. 30.4% of students answered this correctly.

While students understand that half means ‘out of 2 parts’ and hence are able to divide by 2 to answer what is half of 10 kg, students do not understand that in a fraction such as half, the two parts have to be equal.

Class 4	Maths
If a watermelon weighs 10 kg, how much will half the watermelon weigh? _____ kg	Which figure is $\frac{1}{2}$ shaded? Tick (✓) the answer.  A.  B.  C.  D.
Sample Question 27: This is a procedural question that checks if a student is able to recognise that a half is one out of two parts and hence divide by 2. 77.4% of students answered this correctly.	Sample Question 28: This is a conceptual question that checks whether the student recognises that option D is the fraction in which both parts are equal. 41.1% of students answered this correctly.

In language, students could retrieve explicitly stated information if it is asked verbatim from the text they read. However, they find it difficult to make straightforward inferences from explicit information, or interpret and integrate ideas and information given in the passage/story.

Passages used to check comprehension were unseen passages that the students would not have had a chance to be familiar with due to them being absent in the textbook. The passages were graded in terms of difficulty and included short stories, short descriptive texts and authentic material such as understanding the information in a bottle label and a notice put up in school. In class 4 and 6, 3 passages were used and in class 8, 4 passages were used. In all the classes, a PIRLS released passage was included. In class 4, a passage was read out to the students to check comprehension.

Knowledge versus Understanding

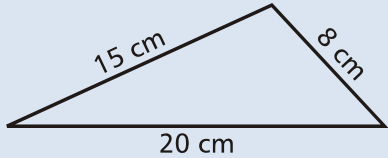
Knowing the facts and doing well on tests of knowledge do not mean that we understand. Bloom (1956) and his colleagues remind us to be specific about how understanding differs from merely accurate knowledge when they recount a famous John Dewey Story:

Almost everyone has had the experience of being unable to answer a question involving recall when the question is stated in one form, and then having little difficulty ...when the question is stated in another form. This is well illustrated by John Dewey's story in which he asked a class, "What would you find if you dug a hole in the earth?" Getting no response, he repeated the question; again he obtained nothing but silence. The teacher chided Dr. Dewey, "You're asking the wrong question." Turning to the class, she asked, "What is the state of the center of the earth?" The class replied in unison, "igneous fusion".

Dewey's story also illustrates the rote recall nature of some knowledge learning. The emphasis on knowledge as involving little more than remembering or recall distinguishes it from conceptions of knowledge that involve understanding or insight, or that are phrased as "really know" or "true knowledge."

Source: Wiggins .G, and McTighe. J, (1998), "Understanding by Design", Page 39.

ANALYSIS OF STUDENT PERFORMANCE IN CLASS 6 QUESTION PAIRS TESTING 'LEARNING WITH UNDERSTANDING'

Rote based /Procedural Questions	% Correct	Understanding /Conceptual questions	% Correct
Write the answer. $\begin{array}{r} 713 \\ \times 24 \\ \hline \end{array}$	47.9%	25 × 18 is more than 24 × 18. How much more? A. 1 B. 18 C. 24 D. 25	21.3%
What is the perimeter of this shape?  _____ cm.	47.9%	A thin wire 20 centimetres long is formed into a rectangle. If the width of this rectangle is 4 centimetres. What is its length? A. 5 centimetres B. 6 centimetres C. 12 centimetres D. 16 centimetres	16.7%
The cost of the pickle is rupees_____.	69.3%	The pickle should not be eaten after the month of _____. (write the month and the year)	26.7%
What made the whole earth shake? A. an earthquake B. an enormous fruit C. the fleeing hares D. a falling tree	51.4%	Write two ways in which the lion tried to make the hare feel better at the end of the story.	13.6%

At class 4 level, students performed slightly better overall in the passage that was read out compared to the one they had to read for themselves. In the international passages, students performed much lower than the international average performance.

In all the classes tested, while students were comfortable in retrieving stated information verbatim from a passage, they showed difficulty in making straightforward inferences from stated information. They also could not interpret information given and integrate ideas when information was not explicitly stated.

Class 8 Language

Passage Excerpt:

An old man who lived in the city had to put up with the nuisance of having boys play cricket in the streets outside his house in the streetlights, at night.

One evening, when the boys were particularly noisy he could not sleep. So he went out to talk to them. He explained that he was very happy whenever he could see or hear boys playing his favourite game, cricket. He said he would give them 25 rupees each week to play in the street at night.

The boys were thrilled.

They were being paid to do something they enjoyed!

At the end of the first weekhe did so.

The second weekaway with only 15 rupees.

The third weekgave them only 10 rupees.

Thefourth weekwould give them 5 rupees each week without fail.

.....The boys then stormed away and never played on the street again.

How much money did he promise to give the boys?

- A. 25 rupees every week
- B. 15 rupees every week
- C. 10 rupees every week
- D. 5 rupees every week

What did the old man really want from the boys?

- A. He wanted them to play cricket on all the days of the week.
- B. He wanted them to practice more and thus improve their game.
- C. He wanted them to stop playing cricket near his house at night.
- D. He wanted them to start earning while doing something they liked.

Sample Question 29: 74.4% of students could retrieve the explicitly stated information correctly

Sample Question 30: 45.7% could integrate the ideas from the information and interpret the information that is not explicitly stated that the old man really wanted the boys to stop playing cricket

Students found filling given words to complete a letter easier than having to actively produce the words by understanding the context in the cloze passage given.

Class 4 Language

Complete the passage using suitable words from the given list.

come, take, grant, suffering, rest

Madam,

As I am _____ from fever, the doctor has advised me to _____. Please _____ me leave for 2 days from today, the 17th of October. I will _____ to school on 19th.

Thank you.

Yours sincerely,

Sheela.

Sample Question 31: 71.0% of students could use the words to complete the passage.

Class 6 Language

Complete the passage using suitable words of your own.

Ganga makes and sells garlands. She buys the flowers from the _____ and she makes them into garlands. She uses flowers of different _____. But she usually makes garlands of small white flowers because they _____ the most. She sells _____ to many people. She sells all the garlands she makes for the day.

Sample Question 32: 38.8% could use appropriate words on their own to complete the passage.

RESTRICTED AND OPEN CLOZE* ITEMS

Sample A: Class 4 – Restricted Cloze Passage

(Overall % correct - 56.1%)

16 नीचे लिखे शब्दों में से उचित शब्द का प्रयोग कर इस पत्र को पूरा करो।

उपस्थित, आया, आराम, छुट्टी, नहीं

महोदया,

मुझे बुखार आया है इसलिए डॉक्टर ने मुझे आराम करने की सलाह दी है। कृपया मुझे आज 17 अक्टूबर, से 2 दिन की छुट्टी दें। मैं दिनांक 19 को बिद्यालय में उपस्थित रहूँगी।

धन्यवाद।

आपकी आज्ञाकारी शिषी

मीना

Complete the passage using suitable words from the given list.

[present, have, rest, leave, not]

Madam,

As I have fever, the doctor has advised me to rest. Please grant me leave for 2 days from today, the 17th of October. I will be present in school on 19th.

Thank you.

Yours sincerely,

Sheela.

Sample B: Class 6 – Cloze Passage

(Overall % correct-38.8%)

13 इन पंक्तियों को अपने उचित शब्दों के प्रयोग से पूरा करो।

गंगा फूलों के हार बनाती और बेचती है। वह फूल बाजार से खरीदकर लाती है और उनसे हार बनाती है। वह इसके लिए कई गंगा के फूल इस्तेमाल करती है। लेकिन वह अधिकतर सफेद फूलों से ही हार बनाती है क्योंकि वह सबसे ज्यादा बेचती है। वह कई लोगों को बेचती है। वह दिन में जितने हार बनाती है, वे सारे बेच देती है।

Complete the passage using suitable words of your own.

Ganga makes and sells garlands. She buys the flowers from the market and she makes them into garlands. She uses flowers of different Ganga for this. But she usually makes garlands of small white flowers because they are the most sells. She sells to many people. She sells all the garlands she makes for the day.

Sample C: Class 8 – Cloze Passage

(Overall % correct- 49.8%)

13 इन पंक्तियों को अपने उचित शब्दों से पूरा करो।

गंगा फूलों के हार बनाती और बेचती है। वह फूल आला से खरीदकर लाती है और उनसे हार बनाती है। वह इसके लिए कई पेड़ के फूल इस्तेमाल करती है। लेकिन वह अधिकतर सफेद फूलों से ही हार बनाती है क्योंकि वह सबसे ज्यादा गंगा है। वह आला को कई लोगों को बेचती है। वह दिन में जितने हार बनाती है, वे सारे बेच देती है।

Complete the passage using suitable words of your own.

Ganga makes and sells garlands. She buys the flowers from the garland and she makes them into garlands. She uses flowers of different tree for this. But she usually makes garlands of small white flowers because they are the most makes garland. She sells garland to to many people. She sells all the garlands she makes for the day.

Sample D: Class 6 - Cloze Passage

13 इन पंक्तियों को अपने उचित शब्दों के प्रयोग से पूरा करो।

गंगा फूलों के हार बनाती और बेचती है। वह फूल बाजार से खरीदकर लाती है और उनसे हार बनाती है। वह इसके लिए कई रंग के फूल इस्तेमाल करती है। लेकिन वह अधिकतर सफेद फूलों से ही हार बनाती है क्योंकि वह सबसे ज्यादा सुन्दर है। वह हूँट कई लोगों को बेचती है। वह दिन में जितने हार बनाती है, वे सारे बेच देती है।

Complete the passage using suitable words of your own.

Ganga makes and sells garlands. She buys the flowers from the market and she makes them into garlands. She uses flowers of different colours for this. But she usually makes garlands of small white flowers because they are the most beautiful. She sells them to many people. She sells all the garlands she makes for the day.

Findings: Students do very well when they are asked to choose words from a list to fit into the context. At the same time, they struggle when asked to write their own words even when the context is quite clear and the possibilities are limited. In both class 6 & 8, the choice of words indicate that students are unable to come up with exact matches that are appropriate to the context. The span or number of words that they can evaluate for context seems to be the crucial factor. In Sample C, for instance, the student has associated the word garland with flowers and written that. The fact that you cannot buy flowers from a garland is missed. The students who have done better seem to be able to evaluate the fitness of the word to a larger set of words or context.

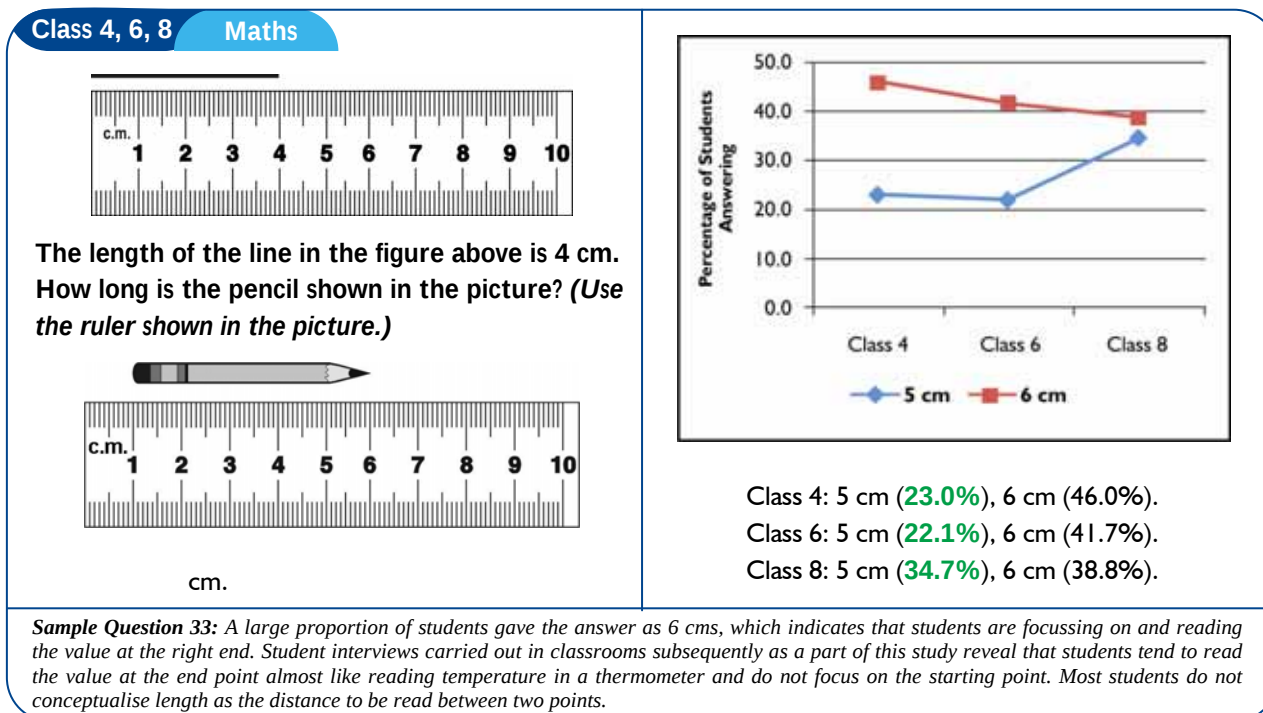
*A **cloze test** is an item consisting of a portion of text with certain words removed, where the participant is asked to replace the missing words. In what we are calling the Restricted Cloze the deleted words are provided separately.

3. Practical competencies such as measurement, problem solving, writing a few sentences that are meaningful and grammatically correct are often among the most difficult questions.

Many competencies that are of practical importance in the real world, although in the curriculum are not acquired by the students. These include understanding measurement, problem solving, writing a few correct sentences, etc.

Measurement of length is often reported in different EI studies (Student Learning in Metros, 2006; Annual Status of Student Learning in Bhutan, 2008) as something that students are finding difficult to acquire. This could be related to the way in which measurement is taught in the Indian curriculum.

Students in the government schools in this study found the understanding of measurement units as well as carrying out measurement of length, weight and volume difficult.



While EI studies in private schools (Student Learning in Metros, 2006; Video Misconception Series) have revealed that students focus on starting and ending point but tend to read the points on the scale, rather than counting the distance between 2 points. In government schools students tend to read off the scale at the end point. The understanding of the process of measurement of length may probably be confounded by the way in which measurement of other quantities are practised in real life. For example, measurement of weight is read as indicated by a needle in a weighing scale, or how temperature is measured by the reading to which the mercury has risen in a thermometer.

Class 4 Maths How many millimetres are in a meter? _____	Class 6 Maths A mango weighs $\frac{1}{2}$ kg. How many grams is that? _____ grams	Class 8 Maths The number of 750 ml bottles that can be filled from 600 L of water is A. 8 B. 80 C. 800 D. 8000												
Sample Question 34: 27.4% of students gave the correct answer as 1000 mm. 14.7% of students gave the answer as 100 mm, while 29.2% did not attempt this question. This is a TIMSS question and internationally 49.0% of students answer this correctly at class 4 level	Sample Question 35: 42.1% of students gave the correct answer as 500, while 17.9% did not attempt the question	Sample Question 36: <table><tr><th colspan="4">% of students choosing</th></tr><tr><th>A</th><th>B</th><th>C</th><th>D</th></tr><tr><td>19.5</td><td>23.4</td><td>29.1</td><td>14.4</td></tr></table> The proportion of students choosing different options seems evenly distributed and indicates that students are attempting a guess to answer this question.	% of students choosing				A	B	C	D	19.5	23.4	29.1	14.4
% of students choosing														
A	B	C	D											
19.5	23.4	29.1	14.4											

Students are not able to tackle problem solving and demonstrate very low ability in such questions.

Class 4

Maths

Rahul is older than Ramesh, and Ramesh is older than Peter. Which statement must be true? Tick (✓) the answer.

- A. Rahul is older than Peter.
- B. Rahul is younger than Peter.
- C. Rahul is the same age as Peter.
- D. We cannot tell who is oldest from the information

Class 8

Maths

Two views of the same cube are shown here

If the cube has different symbols on each face, what symbol will be there on the face opposite ✕?

- A. ★
- B. ✿
- C. ★
- D. Cannot be determined

Sample Question 37: 39.5% could answer this correctly while internationally 63.0% of students could do so

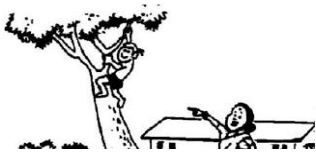
Sample Question 38:

% of students choosing			
A	B	C	D
20.1	25.8	24.4	19.3

The proportion of students choosing different options seems evenly distributed and indicates that students are attempting a guess to answer this question.

Writing seems to be the most difficult skill in language that is seen to be achieved only by the very high performers in the class. Questions on writing ranged from checking skills as simple as spellings, reorganizing a jumbled sentence or sentences, writing a single to a few meaningful sentences for a picture or a topic, completing a cloze passage, completing a dialogue between a shopkeeper and a girl who goes there to buy pencils, and completing a miniature short story by adding 2 sentences from the word clues given.

As explained in chapter 4, in detail, the scale anchoring process (which identifies what students at different ability levels know and are able to do) points out that the key differentiating skill of students who are at the top of the class at 90th percentile and at 75th percentile are those who are able to handle the questions that require writing as described above.

Class 4 Language What could Kumar's mother be telling him in the picture below? Write in a sentence.  _____ _____ _____	Sample Question 39: This is a question that only students in the 90th percentile of class 4 got correct. Students in the 75th percentile in class 6 and the 50th percentile in class 8 could get this correct.
Class 6 Language Complete the following conversation between Leela and the shopkeeper using your own words / sentences. Leela: Do you have pencils available in your shop? Shopkeeper: _____ Leela: I want one pencil. Shopkeeper: _____ Leela: How much does this pencil cost? Shopkeeper: _____ Leela: Here, take 4 rupees. Shopkeeper: _____ Leela: No, I do not want anything else. Shopkeeper: Ok. Come whenever you need to buy pencils.	Sample Question 40: This is a question that only students in the 90th percentile of class 6 and students in the 75th percentile in class 8 got correct.

Rearrange the words to construct a proper sentence.

heavy rains / flooding / several / lashed / low-lying areas / the city / for 72 hours.

Sample Question 41: This is a question that only students in the 90th percentile of class 8 got correct. In class 6, even 90th percentile students could not do this question.

4. Students are coping, as they are possibly learning from daily life experiences than through a formal education in school.

Surprisingly while students could count fractional quantities given in real life situations, fewer could add two fractional quantities in a straightforward addition problem. The straightforward computation of adding 2 mixed fractions can be obtained only through formal education, while understanding 2½ kg and 1½ kg and adding the quantities in real life contexts can be got by learning that need not happen in a school.

Write the answer.

$$2\frac{1}{2} + 1\frac{1}{2} = \underline{\hspace{2cm}}$$

Sample Question 42: 20.5% answered correctly the computation of mixed fractions.

Ravi bought 2½ kg rice from one shop and 1½ kg rice from another shop and put it in his bag. What is the total weight of rice he bought?

_____ Kg

Sample Question 43: 34.6% answered correctly the computation of mixed fractions by using the familiar contexts given.

Similarly while students could identify the least priced shirt from among 2 digit and 3 digit values, they found ordering single digit numbers difficult. This also shows that they are learning in real life contexts outside the school environment but not necessarily connecting the same to the learning in the classroom.

In the list below, tick (✓) the number that is greater than 5 but less than 8.

4	7	9	2
---	---	---	---

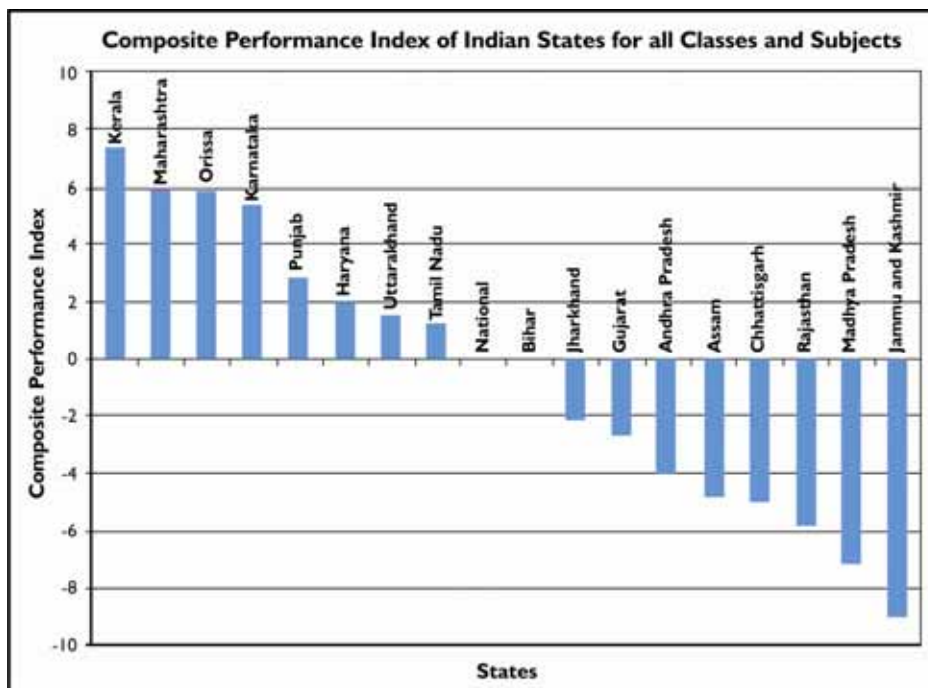
Sample Question 44: 48.1% answered this question correctly

Which of these shirts costs the least? Tick (✓) the answer.



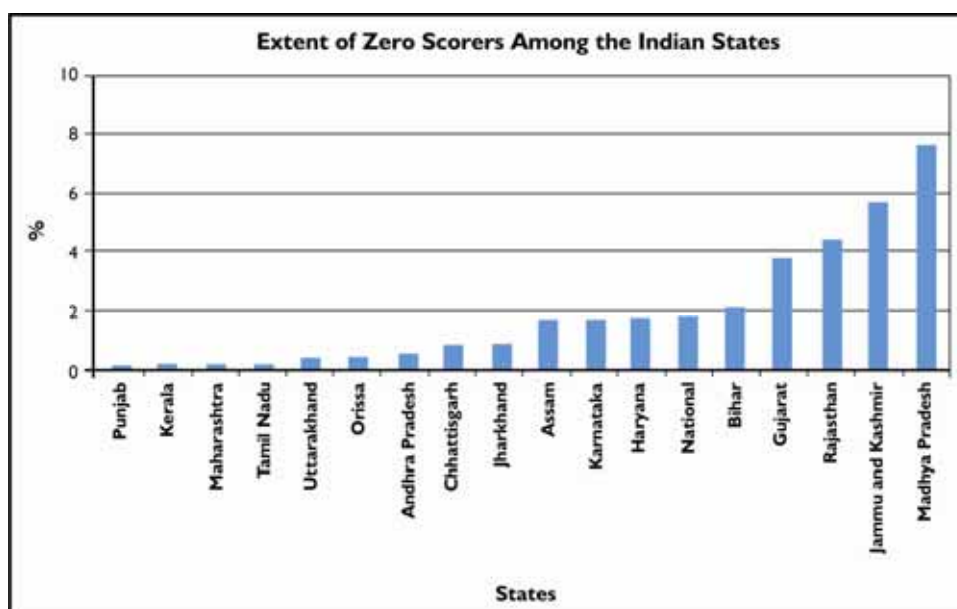
Sample Question 45: 77.1% answered this question correctly

5. Kerala, Maharashtra, Orissa and Karnataka clearly show an overall better performance than the national average. Jammu and Kashmir, Madhya Pradesh and Rajasthan were among the states that ranked among the bottom three overall and also have more students not scoring in the tests, indicating that these students are being left behind in these states.



Graph 2.1: Composite performance index of schools of different states

Based on the relative performance/rank of the states in different classes and subjects, an attempt has been made to consolidate the performance of 17 states. The union territory of Chandigarh, although tested, has not been included in this comparison. Delhi, though a state, is also not included in this comparison as only class 4 was tested here. A high correlation of 0.85 was observed in the performance across the classes and subjects tested in each state, which means if a state does well in a class or subject then it more or less tends to do well in other classes and subjects too.



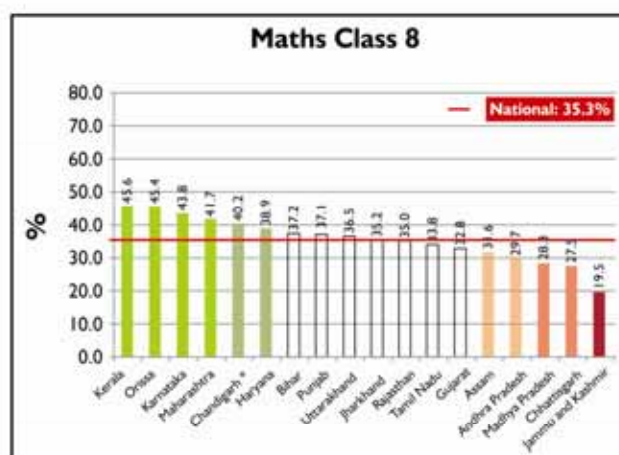
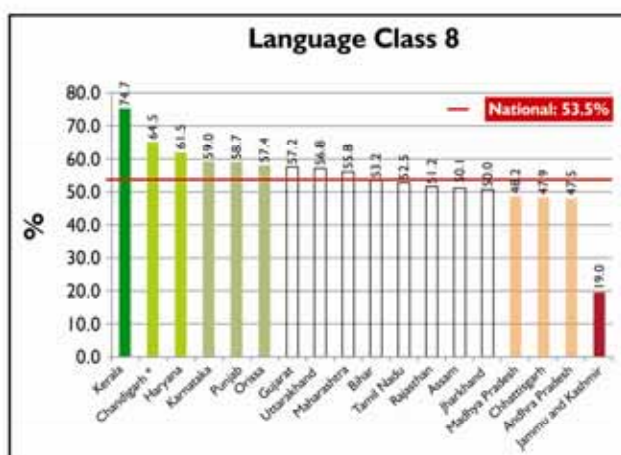
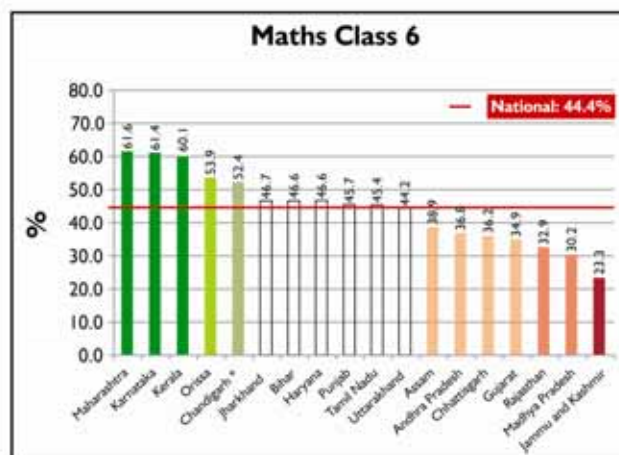
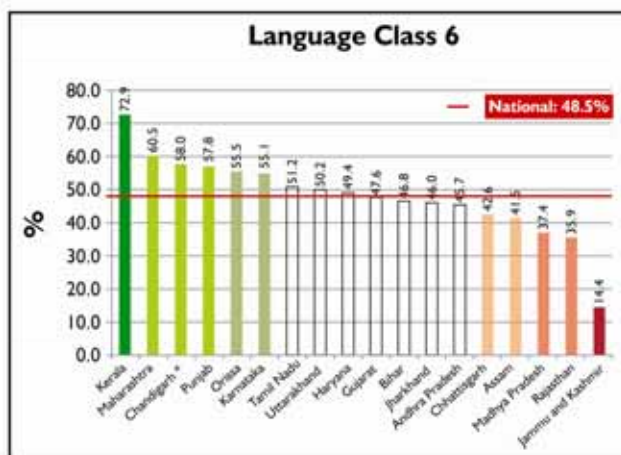
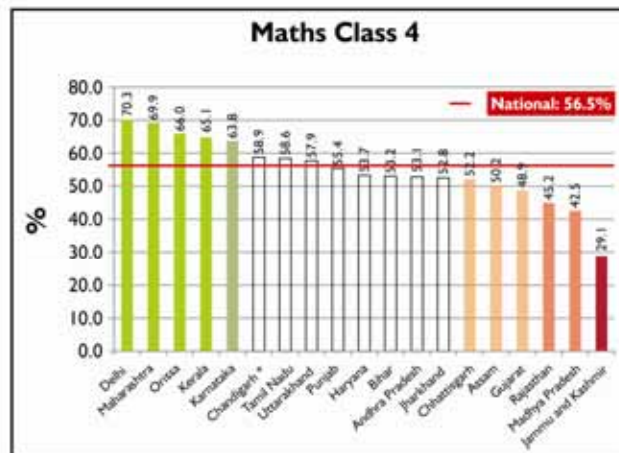
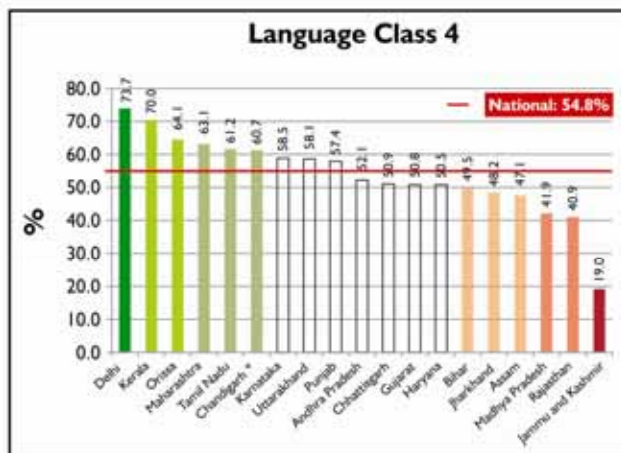
Graph 2.2: The total number of students scoring zero in the classes and subjects tested

Andhra Pradesh, Assam, Chhattisgarh and Jharkhand, although below the national average, have fewer students not scoring in the test compared to better performing states such as Haryana and Karnataka, indicating that while the states' efforts could be addressing the lowest ability students, it does not provide adequate support for overall improvement in all students. While Bihar performs closer to the national average, Gujarat is lower than the national average and also has more students not scoring in the tests.

The performance of each state in the classes and subjects compared to the national average was further checked to identify statistical significance using student's t tests. It was felt that as the samples may show statistical significance due to their large size. Hence, for a more robust interpretation, the relative magnitude /effect size of the observed differences were calculated using Cohen's d. In other words, Cohen's d helps us determine if the differences observed were substantively meaningful.

Table 2.1

Substantive Meaning of the Observed Performance of the States for all Classes and Subjects



The mean difference significant at 0.05 level. Effect size >0.2 – Small(S); >0.5 – Medium (M), >0.8 – Large

Note:- * Union Territory

Difference Above National Average	
	Statistically Significant and Large Effect
	Statistically Significant and Medium Effect
	Statistically Significant and Small Effect

Difference Below National Average	
	Statistically Significant and Large Effect
	Statistically Significant and Medium Effect
	Statistically Significant and Small Effect

Kerala, Maharashtra, Orissa and Karnataka are clearly doing well in almost all classes and subjects tested. However, Maharashtra has to focus on Language in class 8 and Karnataka has to do so in class 4.

Punjab starts to catch up from class 6 in language and needs attention in maths overall. Haryana starts to catch up in class 8 in both language and maths but needs to focus on lower classes.

Although Uttarakhand is above the national average in most classes tested, the observed difference is practically not substantive enough. Tamil Nadu does well in language in class 4, but starts performing lower than the national average in class 8. Similarly, Chhattisgarh is similar to national average in language in class 4, but is substantively lower than the national average in all other classes and subjects.

Bihar and Jharkhand show a substantively low performance compared to national average in class 4 language and are below national average in other classes and subjects. Gujarat is substantively low in maths in class 4 and 6.

Andhra Pradesh and Assam are below the national average in all classes and subjects with Andhra Pradesh substantively lower from class 8 in language and class 6 in maths, while Assam is substantively lower in all classes except in class 8 language. Rajasthan is substantively much lower in class 4 and 6 in language and maths, while Madhya Pradesh and Jammu Kashmir are substantively low in all classes and subjects.

6. On common questions used from an extensive study for private schools in India catering to the elite and upper middle class, students in government schools showed a much lower performance. The purpose of the comparison was not to compare the 2 school systems but rather to identify what is the highest level of achievement that students of that age group could potentially reach.

Class 8
Maths

School Type	Performance on Common Questions (%)
India Public	19.0
India Private	73.5

Which rational number does the pointer on the number line show?

A. -1.2
B. -0.4
C. 1.2
D. -0.8

Sample Question 46: 19.0% of students from Indian public schools answered this correctly while 73.5% answered correctly in Indian private schools

A lot of opinion may exist on whether it is okay in the first place to do a comparison of the Public school students with the private school students. It is well known that elite private schools in India cater to students who are better advantaged on a number of input factors – socio-economic factors such as prevalence in urban areas, educated parents, aspiration levels and school factors related to basic facilities and teacher factors related to accountability and ability.

When parents are possibly not first generation learners and are from an economically advantaged middle class, the importance that could be accorded by the family to learning and the tendency of the parents to prime their wards towards better learning is generally higher (Malcolm Gladwell, 2009). However, teachers have a higher impact on student learning as teachers in private schools have lower teacher absence and higher levels of teaching activity and this produces higher test scores after controlling for observable family and school characteristics (Karthik Muralidharan and Michael Kremer, 2006).

Overall, 4 questions in class 4, 1 in class 6 and 2 in class 4 were used from Educational Initiatives' diagnostic test ASSET⁸. In each of the common questions, the performance of students in public schools was much lower and did not exceed half the performance exhibited by private school students. This indicates the presence of 2 disparate populations in public and private schools and the extent to which we need to close the gap to make available quality learning to all children in India.

⁸ ASSET is a diagnostic achievement test that is offered to private school students across India and internationally. 4 lakh students take ASSET year on year.

The graph alongside shows the number of visits to a website during six months. Answer the question based on it.

Which months recorded the minimum and maximum visits?

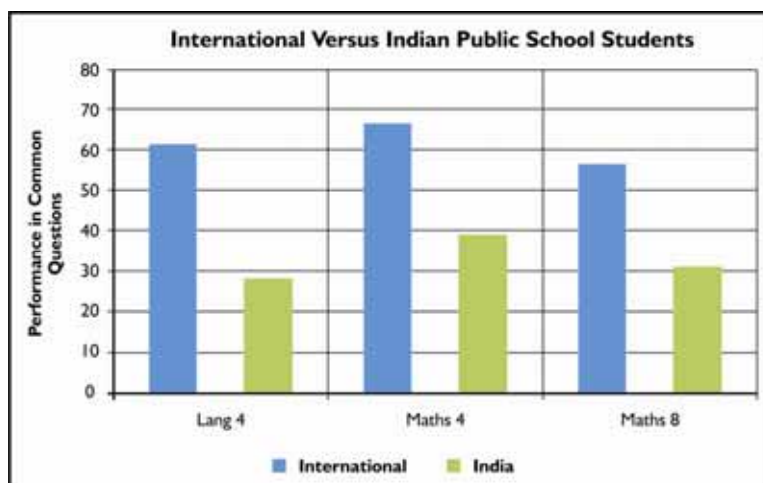
- A. Minimum: March, Maximum: July
- B. Minimum: March, Maximum: August
- C. Minimum: August, Maximum: July
- D. Minimum: April, Maximum: May



Sample Question 47: 48.1% answered this question correctly.

7. The level of learning of Indian students in government schools in class 4 and class 8 tested is much lower than the international average as represented by studies like Trends in International Maths and Science Study (TIMSS) and Progress in International Reading and Literacy Study (PIRLS).

In order to obtain a comparative picture of how our students learn vis-a-vis students internationally, a few questions were carefully selected from the international tests and included in the student learning study. The results show that Indian students lag behind their international counterparts in both language and maths.



Graph 2.3: The percentage of international and Indian public school students scoring zero in the classes and subjects tested.

Which of these has the same value as 342? Tick (✓) the answer.

- A. $3000 + 400 + 2$
- B. $300 + 40 + 2$
- C. $30 + 4 + 2$
- D. $3 + 4 + 2$

Sample Question 48: Internationally 87.0% of students answered this correctly, while 43.7% of students from Indian public schools could get this correct.

Passage Excerpt:

When Lakhan discovered that he had mice in his house, it did not bother him much at first. But the mice multiplied. They began to bother him. They kept on multiplying and finally there came a time when even he could stand it no longer.

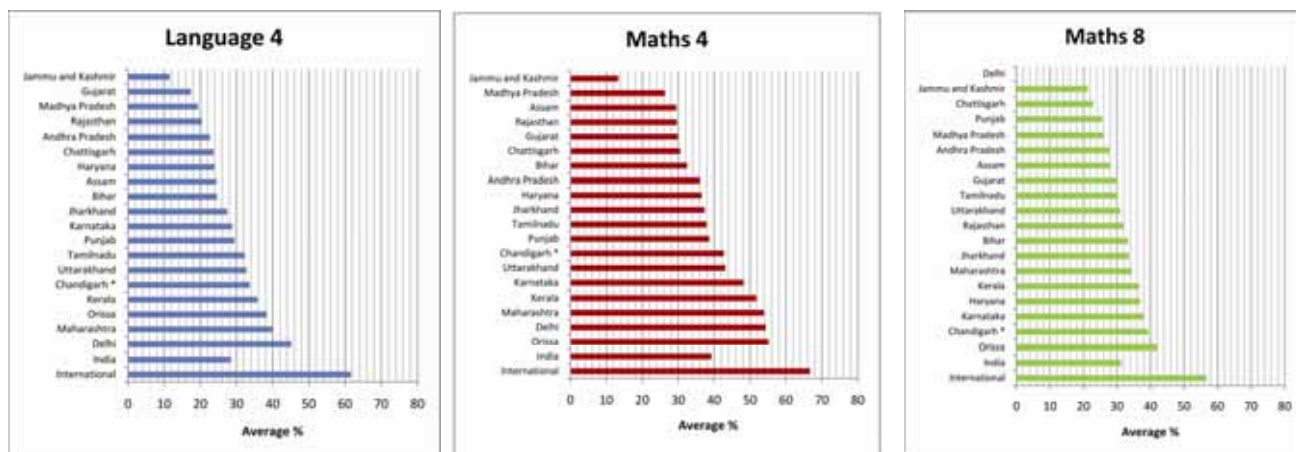
Why did Lakhan want to get rid of the mice?

- A. He had always hated mice.
- B. There were too many of them.
- C. They laughed too loudly.
- D. They ate all his cheese.

Question involves retrieving stated information in the text and making straight forward inferences from it.

Sample Question 49: Internationally, 79.0% of students answered this correctly, while 41.8% of students from Indian public schools could get this correct.

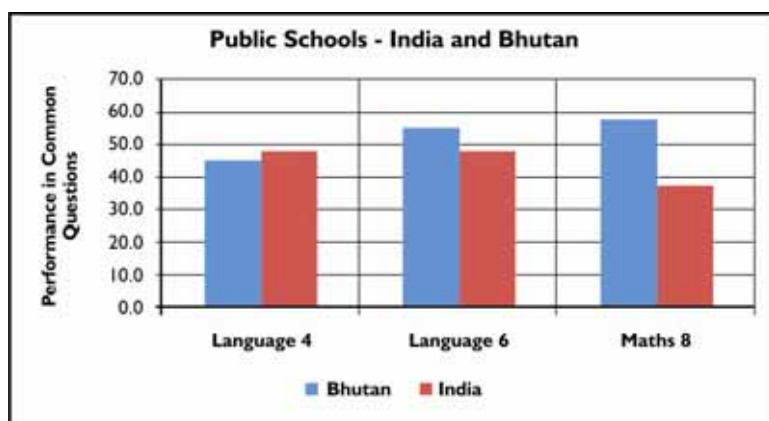
The same trend of performance lower than the international average is seen in all the Indian states that participated in the study. This indicates that the prevalence of low learning levels compared to the international average remains consistent across states in this category of schools. It is not that some states perform exceptionally low compared to others.



Graph 2.4: The percentage of students answering the common questions from international tests correctly

8. A comparison of government schools in neighbouring Bhutan, relevant from the point of a school system facing similar socio-economic, management and accountability factors reveals that the level of learning in the public schools of India is slightly higher in language⁹ at Class 4 level. This could possibly be due to the advantage of learning in the mother tongue compared to all of Bhutan's students who learn in English as the medium of instruction from Class 1. However, any learning gains Indian students have due to learning in the mother tongue seems to be lost as students move up to higher classes 6 and 8, with Bhutanese students doing better than their Indian counterparts in both language and maths

This comparison facilitates us to look at other educational systems from neighbouring countries that are also beset by similar development/socio-economic problems that may impact the learning quality in government schools. The results show that students in India do significantly better in language at Class 4. Research in student learning indicates that students come to classrooms equipped with a fairly decent vocabulary in their home language and this enables them to get a head start if the learning in school is in the same language. Students in Indian public schools have this advantage of learning in the mother tongue compared to all of Bhutan's students where students learn in English as the medium of instruction from Class 1. As the students move up to higher classes, it does not seem to be important to have the mother tongue as the medium of instruction as could be seen from the Annual Status of Student Learning study (ASSL 2008). SLS reveals a similar pattern where at the Class 6 level Bhutanese students start doing significantly better in language although they are learning in English which is not their mother tongue. In Maths 8, Indian students do not do as well as their Bhutanese counterparts.



Graph 2.5: The performance in common questions

Class 6 Language

Passage Excerpt:

The lamb said, "Please, please don't eat me yet. My stomach is full of grass. If you wait a while, I will taste much better.

The wolf sat down and waited. After a while, the lamb said, "If you allow me to dance, the grass in my stomach will be digested faster." Again the wolf agreed. While the lamb was dancing, she said, "Please take the bell from around my neck. If you ring it as hard as you can, I will be able to dance even faster."

The lamb said that _____.

- A. she had not eaten yet
- B. her stomach was full of grass
- C. she was a fast dancer
- D. the bell was very hard

Sample Question 50: 76.4% of Bhutanese students answered this correctly, while 55.1% of students from Indian public schools could get this correct.

⁹ Language in Bhutan refers to testing in English proficiency and language in India refers to testing the regional language proficiency.

2.2 Subject Wise Misconceptions and Common Errors

Misconceptions abound even in very basic concepts like the value of a digit in a number, the shapes of figures such as a square and triangle, decimals, fractions and their representation, and much more...

Students develop a cognitive understanding of the world around them through interactions based on their daily experiences. Teachers and schools help build this understanding. 'Misconceptions' are concepts that students acquire which are not in line with or do not match the conventional expert view in that topic. These result in cognitive gaps in their understanding. Students seem to fill these gaps by formulating their own notions, attributing meanings, and by drawing conclusions that might appear as logical to them. The resulting misunderstandings or alternative concepts formulated by the students, if not challenged, interfere with subsequent learning. As a consequence, students experience difficulty in understanding and internalising higher concepts in that area. Generally, student misconceptions persist until students recognise that their understanding is flawed.

Identifying the exact nature of student misconceptions is difficult through regular classroom interactions. Any question in an assessment test that attempts to identify the misconceptions of the students must force the students to actively use their conceptual understanding. A detailed understanding of these misconceptions through large scale diagnostic assessments provide the teacher with a starting point to explore these in the classroom and eventually help to build correct notions or conceptual understanding.

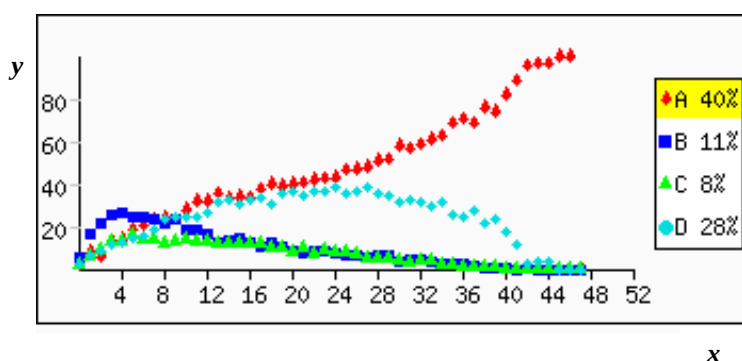
Detailed analysis of the answers given by students in each question in both maths and language reveal that students' understandings of the concepts are often faulty as they seem to harbour many mistaken notions. They also seem to commonly make errors in answering some questions. For example, students harbour wrong notions such as:

- The length of an object is usually the unit on a scale where the measurement terminates. The starting point is not of significance.
- The bigger the number, the higher the place value of all its digits too.
- The bigger number is the one with more digits and the decimal point is not of significance in deciding the same.
- The first day of each month starts on the same day.
- A triangle is a figure with a vertex at the top.
- A square can be recognised as a square only in an upright orientation (with one side horizontal) and not when it is tilted.

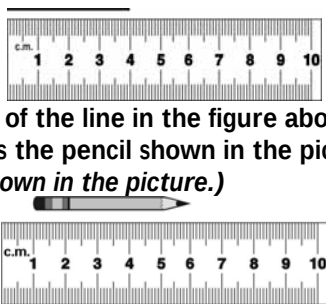
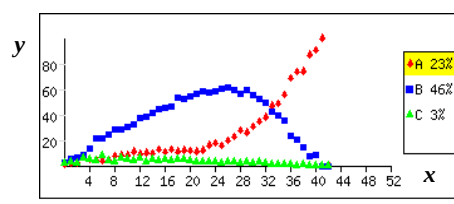
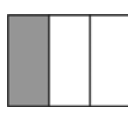
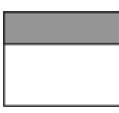
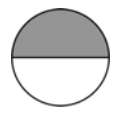
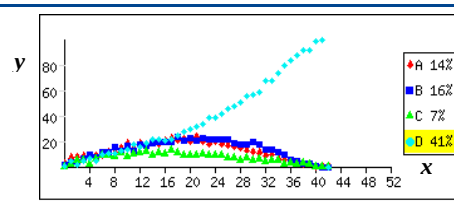
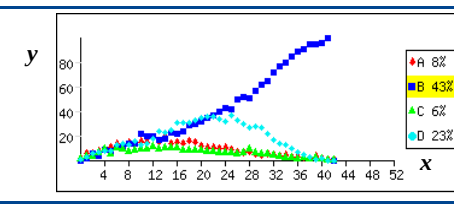
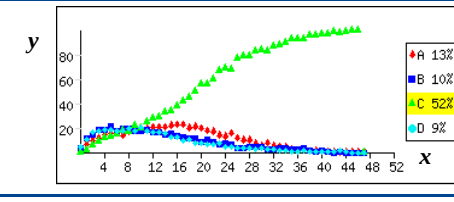
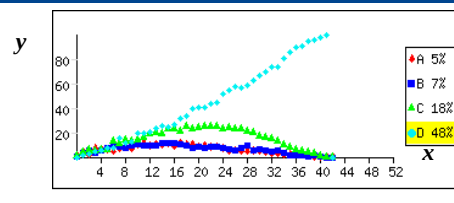
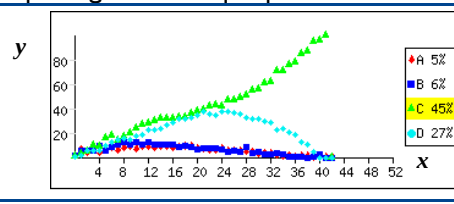
and much more....

How to read the graph:

These are the item response curves for the question. The X axis is the total score in the paper and the Y axis is the percentage of students. The graphs show what percentage of students scoring different total scores in each class, chose which answer options for the question. The legend box shows the total percentage of students who chose each option in the paper. In this graph, the percentage of students choosing option A increases as one moves from a low total score to a high total score, while percentage of students choosing option D increases upto a certain total score and then starts reducing.



2.2.1 Math Misconception and Common errors

S. No.	Class	Q. No.	Question	Graphs
1	4	42	 The length of the line in the figure above is 4 cm. How long is the pencil shown in the picture? (Use the ruler shown in the picture.) _____ cm.	 A: 5 cm, B: 6 cm, C: 4 cm
46.0% of students think that the length of an object is usually the number on the scale corresponding to the end point of the object irrespective of starting point of the object on the scale.				
2	4	23	Which figure is $\frac{1}{2}$ shaded? Tick (✓) the answer.  A.  B.  C.  D.	
Students choosing option A (around 14%) do not understand that the figure should be equally divided into two parts then only can it be $\frac{1}{2}$ shaded. Students choosing option B (around 16%) do not understand that one out of two parts is $\frac{1}{2}$. They seem to interpret the figure in option B as 1 part shaded and 2 parts not shaded and hence think that the shaded part can be represented as $\frac{1}{2}$.				
3	4	26	Which of these has the same value as 342? Tick (✓) the answer? A. $3000 + 400 + 2$ B. $300 + 40 + 2$ C. $30 + 4 + 2$ D. $3 + 4 + 2$	
	6	20	Which of the following is true? A. $69 = 6 + 9$ B. $69 = 6 + 90$ C. $69 = 60 + 9$ D. $69 = 60 + 90$	
23.8% of Class 4 students think that $342 = 3 + 4 + 2$, while 13.7% of Class 6 students think that $69 = 6 + 9$, indicating the lack of understanding of place value of digits in a number. They seem to think of a number as digits joined together rather like beads in a string.				
4	4	10	Which figure is divided into four EQUAL parts? Tick (✓) the answer.  A.  B.  C.  D.	
Only about 48.4% of the students could identify the shape divided into 4 equal parts among the given shapes. 18.3% of students chose option C indicating that they are looking at equidistant lines and interpreting these as equal parts.				
5	4	20	Which of these is a triangle? Tick (✓) the answer.  A.  B.  C.  D.	
Nearly 1/4th of the class tested do not understand that a triangle is a figure with 3 straight sides. These students are probably familiar only with an equilateral triangle in an upright orientation and hence get confused with option D, not realising that it is a 4 sided figure.				

6	6	45	All of the pupils in a class cut out paper shapes. The teacher picked one out and said, "This is a triangle." Which of these statements MUST be correct? A. The shape has three sides. B. The shape has a right angle. C. The shape has equal sides. D. The shape has equal angles.	
---	---	----	---	--

A little less than $\frac{1}{2}$ (48.2%) of the class 6 students answered correctly and demonstrate the understanding that a triangle is a shape having 3 sides. 10.7% and 13.1% of students respectively think that all the triangles need to have a right angle and all triangle have equal sides.

7	6	33	The area of which of these is different from the other three? A. B. C. D.	
---	---	----	--	--

Only 28.8% chose the correct option C. About 36.1% of class 6 students chose option B probably because it has a different shape than the rest, showing that they do not have an understanding of the concept of area.

8	6	5	In which of these numbers below does 3 have the greatest place value? A. 136 B. 239 C. 301 D. 743	
---	---	---	--	--

About 35.8% of the students think that greater the number, the greater is the place value of all its digits. Some students may also have chosen option C because in the number 301, 3 is the greatest digit.

9	6	32	Which tree is tallest? A. Height of tree (1) : 7 metres B. Height of tree (2) : 4.2 metres C. Height of tree (3) : 3.25 metres D. Height of tree (4) : 2 metres	
---	---	----	--	--

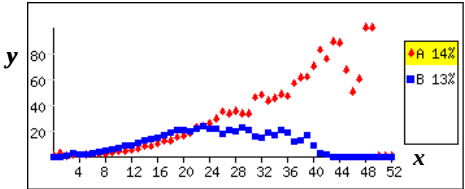
About 40.4% of the students understand the ordering of positive decimal numbers up to 2 decimal places. About 30.8% of students think that 3.25 is greater than 7 and have answered C. They may be having the wrong notion that more the number of digits after a decimal point, larger is the number. It is quite possible that they may have ignored the decimal point and applied the ordering of whole numbers to answer wrongly as C.

10	6	25	Which shape is this? A. a square B. a triangle C. a circle D. a cube	
----	---	----	--	--

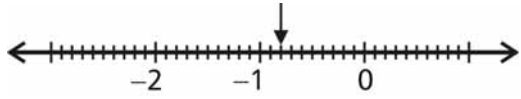
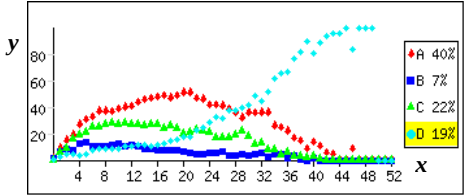
About 40.9% of the students could identify the shape correctly as square. About 28% of students think that a tilted square is a cube.

11	6	27	Look at this calendar. <table><tr><th colspan="7">March 2009</th></tr><tr><td>Sunday</td><td>1</td><td>8</td><td>15</td><td>22</td><td>29</td><td></td></tr><tr><td>Monday</td><td>2</td><td>9</td><td>16</td><td>23</td><td>30</td><td></td></tr><tr><td>Tuesday</td><td>3</td><td>10</td><td>17</td><td>24</td><td>31</td><td></td></tr><tr><td>Wednesday</td><td>4</td><td>11</td><td>18</td><td>25</td><td></td><td></td></tr><tr><td>Thursday</td><td>5</td><td>12</td><td>19</td><td>26</td><td></td><td></td></tr><tr><td>Friday</td><td>6</td><td>13</td><td>20</td><td>27</td><td></td><td></td></tr><tr><td>Saturday</td><td>7</td><td>14</td><td>21</td><td>28</td><td></td><td></td></tr></table> What day will 1st April be?	March 2009							Sunday	1	8	15	22	29		Monday	2	9	16	23	30		Tuesday	3	10	17	24	31		Wednesday	4	11	18	25			Thursday	5	12	19	26			Friday	6	13	20	27			Saturday	7	14	21	28			<table><tr><td>A</td><td>50%</td></tr><tr><td>B</td><td>16%</td></tr></table> A: Wednesday, B: Sunday	A	50%	B	16%
March 2009																																																																
Sunday	1	8	15	22	29																																																											
Monday	2	9	16	23	30																																																											
Tuesday	3	10	17	24	31																																																											
Wednesday	4	11	18	25																																																												
Thursday	5	12	19	26																																																												
Friday	6	13	20	27																																																												
Saturday	7	14	21	28																																																												
A	50%																																																															
B	16%																																																															

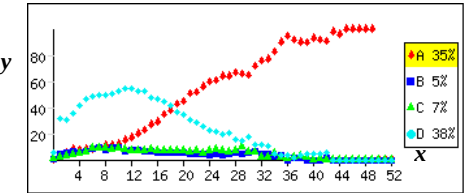
Only 50.5% of the students could correctly reason out that if 31st March falls on Tuesday, 1st of April would be Wednesday. About 15.2% of students wrote Sunday, they either do not understand that the calendar is for March or they think that all months start on a Sunday.

12	8	8	$5 + 4 \times 3 \div 2$ $= 5 + 12 \div 2$ $= 5 + 6$ $= 11$ What is $20 - 8 \div 4 + 7$? _____	 A: 25, B:10
----	---	---	--	--

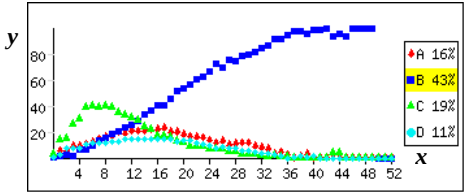
Only about 14.7% could apply BODMAS/PEDMAS rule correctly though it was illustrated in the question by an example. 13.7% of students chose 10 as the answer, indicating that while they may know whole number operations, they do not know the order in which they have to do them in a situation like this.

13	8	18	Which rational number does the pointer on the number line show?  A. -1.2 B. -0.4 C. 1.2 D. -0.8	
----	---	----	--	---

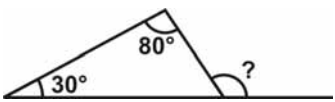
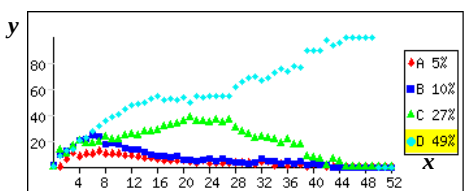
Only about 19.0% of the students chose the correct option D (-0.8). About 40.8% of the students chose -1.2. This suggests that they either do not understand the direction in which the negative numbers increase in a number line, or may not understand that for negative numbers, the number with a smaller value is a bigger number and hence -1 should be to the right of -1.2.

14	8	6	Which of the following gives the largest quotient? A. $478 \div 5$ B. $478 \div 9$ C. $478 \div 7$ D. $478 \div 12$	
----	---	---	---	--

About 38.6% of the students have the wrong notion that the larger the divisor, the larger the quotient when a number is divided. Only about 35.4% of the students chose the correct option A.

15	8	36	Which of these expressions is equivalent to y^3? A. $y + y + y$ B. $y \times y \times y$ C. $3y$ D. $y^2 + y$	
----	---	----	--	---

About 43.5% of the students could understand the exponent notation and choose the correct option B. About 19.6% of the students think that y^3 is equal to 3 times y ($y^3 = 3y$).

16	8	21	In the following figure, what will be the measure of the angle marked ' ? '?  A. 30° B. 80° C. 70° D. 110°	
----	---	----	---	---

About 27.3% of the students chose the wrong option C. They have incorrectly applied the angle sum property which states that the sum of the three interior angles of a triangle is always 180 degrees, in which they have ignored that the angle whose degree is asked is not an interior angle of a triangle. About 49.8% of the students seem to have understood the exterior angle property and chosen the correct answer D.

2.2.2 Language Common Errors and Limitations

S.No.	Class	Q.No.	Question	Graphs
1	4	9	Read the sentences below and tick (✓) the answer to the question. Radha: Every day I get up at 5 o'clock in the morning. Priya: I am still having dreams at that time. Neela: Oh! By that time I finish my prayers! Who gets up first? A. Radha B. Priya C. Neela D. They all get up at the same time.	
41.5% of students chose Radha as the person getting up first. This could be because Radha explicitly mentions the time she wakes up. They could also be judging that 5 o'clock in the morning is quite early based on their every day experience and hence fail to take cognisance of the information subsequently provided by Neela that she finishes her prayers by that time. This could also mean that the students are failing to connect ideas that are implicit, for example, if Neela has to finish her prayers by 5 o'clock in the morning, then she has to wake up much earlier than that.				
2	4	6	Question: _____ tore this paper? Answer: Srikanth tore it. A. How B. Why C. Who D. How many	
About 40.1% of students could match the correct interrogative word that will elicit the answer provided in the dialogue. In vernacular form, the word "why" may also fit in as a standalone question sentence as in "Why was this paper torn?" However, such a 'why' question will not obviously elicit the response that has been provided. Students choosing the wrong options are not connecting the ideas from the information while they read.				
3	4	13	Passage excerpt: But the babies cannot dig their way out of the sand above them. They peep out and make some sound while hatching out of their shells. At that time, the mother, who has been guarding them hears them. She then digs them free. Why does the mother crocodile dig her babies out of the sand? A. The babies cannot come out by themselves. B. She is playing a game with them. C. She is saving them from enemies. D. The babies start fighting with each other.	
In the passage, the reason as to why the mother crocodile dig the babies out of the sand precedes the action of the mother crocodile digging them out. There is also an intervening sentence between the sentences giving the reason and the actual action. This could also have made it difficult for the students to connect the two pieces of information. However, about 40% of the students have managed to do this.				
4	6	28	Passage excerpt: ... Read the information in this wrapper (label) of Star Pickles and answer questions Please keep in a cool and dry place... Which of these would be a suitable place to keep the pickle bottle? A. Outside in the open courtyard B. In the storeroom in the corner C. Next to the stove in the kitchen D. In a shelf in the bathroom.	
The passage states that a cool, dry storage place is required. Nearly 3/4th of the students tested have chosen evenly among all options other than option 4. This indicates that they are guessing and the only option they have been able to rule out is on the basis that a food item would not be stored in a bathroom.				

5	6	23	Passage excerpt: .. A wolf attacked the lamb. The lamb said, “..... don’t eat me yet. If you wait a while, I will taste much better.” ...The wolf ... waited., the lamb said, “If you allow me to dance, ...”.....wolf agreed. While the lamb was dancing, she said, “... take the bell ... ring it as hard The wolfrang it as hard The shepherd heard the bell ringing sent his dogssaved the lamb’s life. Which words below best describe the wolf? A. greedy and clever B. innocent and hungry C. foolish and greedy D. clever and hungry	
---	---	----	---	--

About 34.3% of students are able to read a story and analyse the behaviour of one of the main characters portrayed in it and evaluate it as ‘foolish and greedy’ based on the unfolding of events described. 17.5% of students chose ‘greedy and clever’ and are not able to differentiate the actions of a clever person vis-a-vis a foolish person. About 30% of students are influenced by the fact that the wolf wants to eat the lamb and hence identify the option with the description ‘hungry’ in it.

6	6	21	The lamb asked the wolf to wait because she _____. A. wanted to taste better for the wolf B. wanted to dance for the wolf C. wanted the wolf to get hungry D. wanted time to plan her escape	
---	---	----	--	--

63.0% of students chose the options A, B and C. This suggests that they are going by what the lamb stated to the wolf and are not able to analyse the situation to get to the real intention behind the lamb’s action. These students thus have read the story superficially and do not comprehend enough to come to right conclusions.

7	8	23	Passage Excerpt: An old man who lived in the city had to put up with the nuisance of having boys play cricket in the streets outside his house in the streetlights, at night. One evening, when the boys were particularly noisy he could not sleep. So he went out to talk to them. He explained that he was very happy whenever he could see or hear boys playing his favourite game, cricket. He said he would give them 25 rupees each week to play in the street at night. The boys were thrilled. They were being paid to do something they enjoyed! At the end of the first weekhe did so. The second weekaway with only 15 rupees. The third weekgave them only 10 rupees. Thefourth weekwould give them 5 rupees each week without fail. The boys then stormed away and never played on the street again. What did the old man really want from the boys? A. He wanted them to play cricket on all the days of the week. B. He wanted them to practise more and thus improve their game. C. He wanted them to stop playing cricket near his house at night. D. He wanted them to start earning while doing something they liked.	
---	---	----	--	--

About 45.7% of students could look past the stated facts to understand the real intention behind the old man’s action. The remaining students are not able to use the information provided to analyse the character’s actions and arrive at the right conclusion.

8	8	25	Passage Excerpt: The world's highest mountains are the 14 mountains in Central Asia that rise higher than 8,000 meters above sea level. The highest of them all is Mt. Everest which reaches 8,848 metres. The mountain, which is a part of the great Himalayan range is located on the Tibet and Nepal border. Mount Everest is famous because _____. A. it is a very beautiful mountain B. it is in the great Himalayan range C. it is the highest mountain above sea level D. it is on the Tibet and Nepal border	
---	---	----	---	--

Students find it difficult to answer questions that require understanding something that is implicit in the passage. E.g. the word “famous” is not used with respect to Mt. Everest. But its important features are mentioned. Students have to choose which of these features makes it famous. Additionally, the fact that Mt. Everest is the “highest mountain above sea level” is not stated verbatim in the passage.

Less than a third of the students (32.3%) have been able to answer this question. The most common wrong answer shows another tendency of the students – they choose facts stated verbatim in the passage (D) even when they do not answer the question asked.

9	8	14	Given below are 4 sentences. Choose the right sequence in which these sentences have to be arranged to read meaningfully. - One day, he was immersed in deep meditation and was unaware of his surroundings. - They hurriedly stole the belongings in the hermitage and started running. - At this time, some robbers reached there. - Sage Mandavya was a great sage who lived many years ago. A. 1 – 2 – 3 – 4 B. 1 – 4 – 3 – 2 C. 4 – 1 – 3 – 2 D. 4 – 3 – 1 – 2	
---	---	----	---	--

This question asks the students to order a set of sentences describing a very short story. All sentences have clear cues that provide guidance to their position in the story sequence. These cues refer to time such as “one day”, “at this time”, “many years ago” and some are pronouns like “he” and “they”. Only 38.1% use these to get the sequence correct

2.3 Learning across Classes

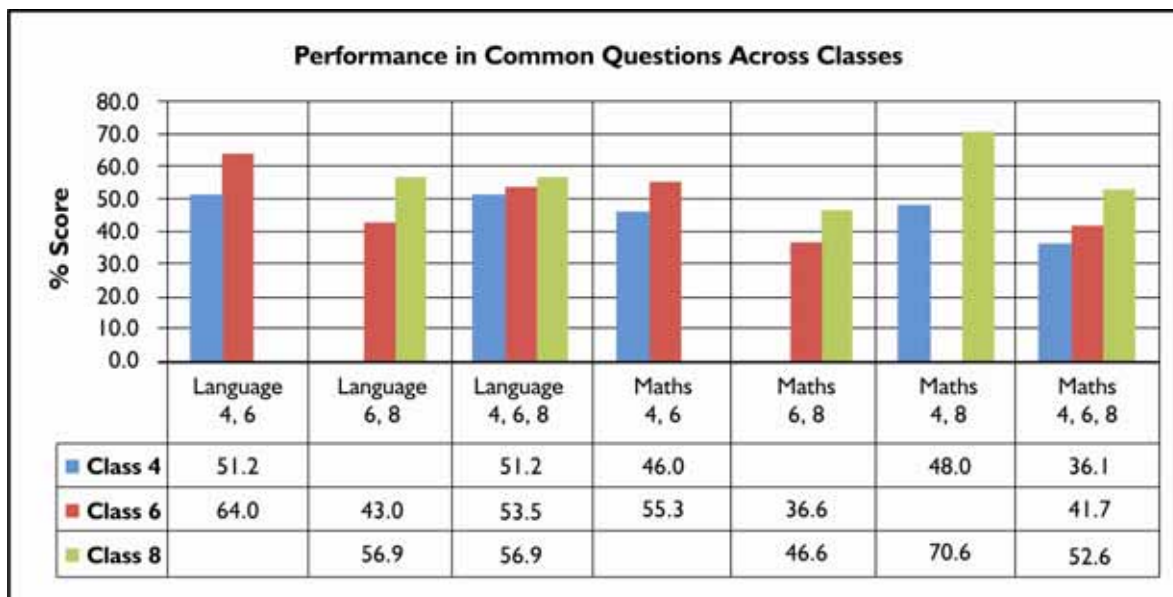
A few questions were used commonly across classes to check if there are learning gains across classes on the same topics. It seems logical and rather obvious that performance would improve as one moves to higher classes on the same topic and reach the peak and then level off. However, does the data also show that? Is the increase substantial and reaching the peak as expected? Are there topics in which students do not do well in higher classes as well?

Table 2.2

Number of Common Questions Across Classes					
Language 4	Language 6	Language 8	Maths 4	Maths 6	Maths 8
6	6		12	12	
	15	15		11	11
			3		3
4	4	4	3	3	3

The number of questions that were common among the papers in the study is given above. Analysis of these common questions across classes tries to find the answers to the questions stated above.

I. Student performance clearly increases in the common questions as students move from class 4 to 6 to 8 across language and maths. However, as can be seen from the graph below, the extent of improvement is often slightly incremental and not a large jump as one would expect.

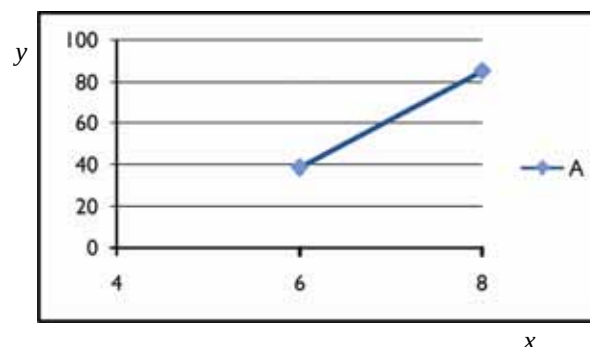


Graph 2.6: Student across class 4,6 and 8 answering common questions correctly

Class 6,8 Language

Complete the passage using suitable words of your own.

Ganga makes and sells garlands. She buys the flowers from the _____ and she makes them into garlands. She uses flowers of different _____. But she usually makes garlands of small white flowers because they _____ the most. She sells _____ to many people. She sells all the garlands she makes for the day.



A: Full Credit(code 01) + Partial Credit (codes 11+21+31)

Code Description

01 - All four words written correctly

First blank – market/bazaar/ Other correct variation.

Second blank – kinds/types/colours/ Other correct variation.

Third blank – sell/ Other correct variation.

Fourth blank –garlands/ Other correct variation.

11 - Three words written correctly in correct place

21 - Two words written correctly in correct place

31 - One word written correctly in correct place

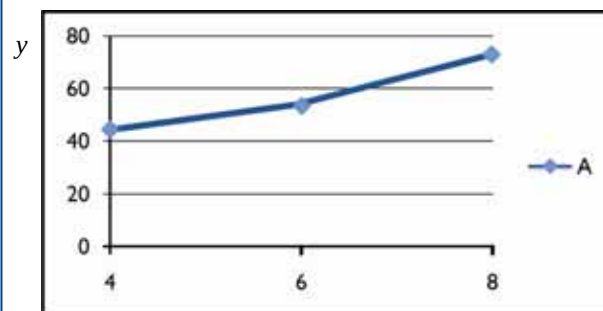
Sample Question 51: This is a cloze test in which students have to complete the passage by using words appropriate to the context.

x –classes tested; y – percentage of students who answered correctly.

It is also seen that the nature of increase is different for different types of questions. For example in sample question 51, which is a cloze item that the student is expected to fill using appropriate words from his own vocabulary, there is a surprising jump with scores almost doubling between class 6 and 8. Similarly, in sample question 52, there is a fair amount of increase in scores between classes 4 to 8. However, questions where large increases in scores are observed are few, and with most of the common questions the increase is slightly incremental with not much change in scores between classes 4, 6 and 8

Write the appropriate number in the empty box.

$$19 + \boxed{} = 32$$



A : 13 (Correct Answer)

Sample Question 52: This question checks the concept of addition where the sum of 2 numbers is equal to a third number. However to get the question correct, the student needs to subtract from the sum, the number that is known, to arrive at the unknown number.

x – classes tested; y – percentage of students who answered correctly.

II. Concepts which should be learnt in lower classes are being understood much later, and even then most of the students are still not acquiring these lower class competencies.

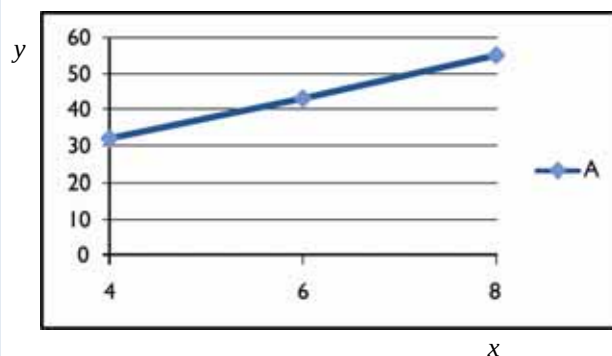
One would expect a class 3 concept to be understood by many students in class 4 and, progressively, for all students to get it right by class 6. If this is the case, there should be a sharp increase in scores between class 4 and 6 while between class 6 and 8 there should not be any marked increase in scores. However, in practise this does not happen.

For example, in the sample question 53 below, the increase from class 4 to class 6 and then class 6 to class 8 is comparable and in sample question 54 which is a class 3 concept asked in class 6 and 8, the increase from class 6 to class 8 is only slightly incremental. These suggest that the performance in the lowest class in which the question is asked is low and not many more students acquire these in higher classes. Even in class 8, the number of students acquiring lower class competencies is low, with most students not acquiring these. For example, in sample question 53, 55.2% of students getting it correct still shows low performance for class 8.

Fill in a suitable word to complete the sentence.

Usually, I remember to bring my lunch box to school.

Whenever I _____ to bring my lunch box I have to go hungry.



A. Forget (Correct Answer)

Sample Question 53: This question checks for usage of opposite words appropriate to a context.

x – classes tested; y – percentage of students who answered correctly.

III. Many misconceptions that students have in lower classes still continue in higher classes and in some cases even become stronger.

Often, logically one would expect that misconceptions are cleared when students learn advanced concepts in the same topics as they progress to higher classes. However, the results suggest that most students who hold a certain type of misconception continue to hold the same or a variant of it as they move to higher classes.

When we analyse the percentage of students choosing the most common wrong answer (which indicates a strong misconception among students), one finds that this percentage is fairly consistent as students move to higher classes. For example, in sample question 55 below, the percentage of students choosing B remains around 16% in all the 3 classes tested. This shows that the number of students who interpret the figure in option B as 1 part shaded and 2 parts not shaded and

hence think that the shaded part can be represented as $\frac{1}{2}$ probably continue to hold the same misconception as they move to higher classes.

Class 4,6,8 Maths

Which figure is $\frac{1}{2}$ shaded?



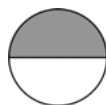
A.



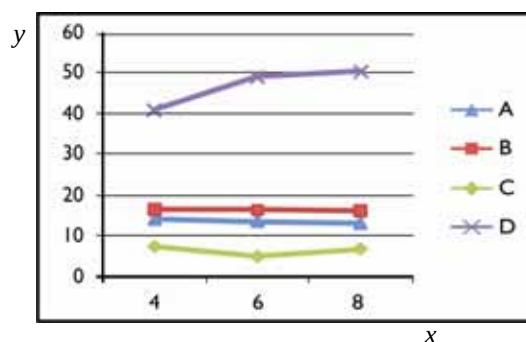
B.



C.



D.



Sample Question 54: This is a question that checks for the concept of $\frac{1}{2}$. To get this correct, student needs to understand equality of parts in a given shape.

x –classes tested; y – percentage of students who answered correctly.

Similarly, in sample question 55, about 36-38% of students chose option B in classes 6 and 8, probably because it has a different shape than the rest and hence its area is different from the other three. In sample question 56, the number of students who recognise a tilted square as a cube actually increases from class 6 to 8. This is probably because they may not be identifying a square based on its properties and hence, are not able to identify one in a different orientation. They also could be working more with 3D shapes such as cubes in higher classes and hence may be confusing a tilted square with a 2D representation of one face of a cube.

Class 6,8 Maths

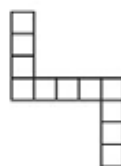
The area of which of these is different from the other three?



A.



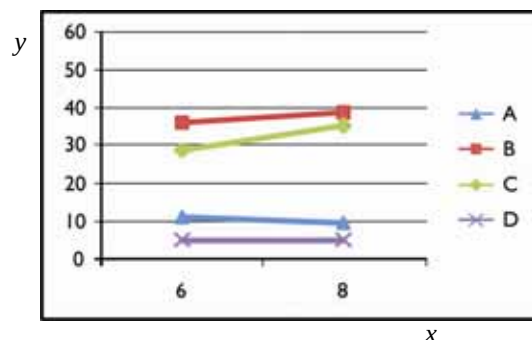
B.



C.



D.

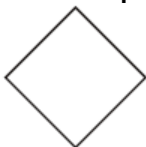


Sample Question 55: This is a question that tests understanding of the concept of area.

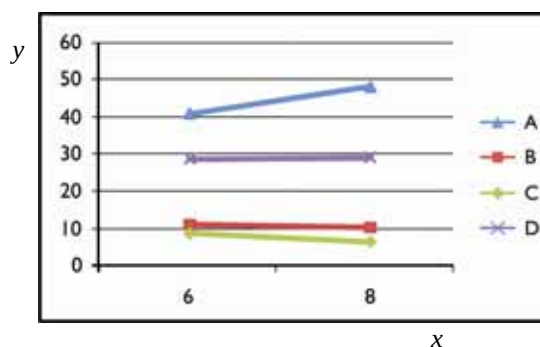
x –classes tested; y – percentage of students who answered correctly.

Class 6,8 Maths

Which shape is this?



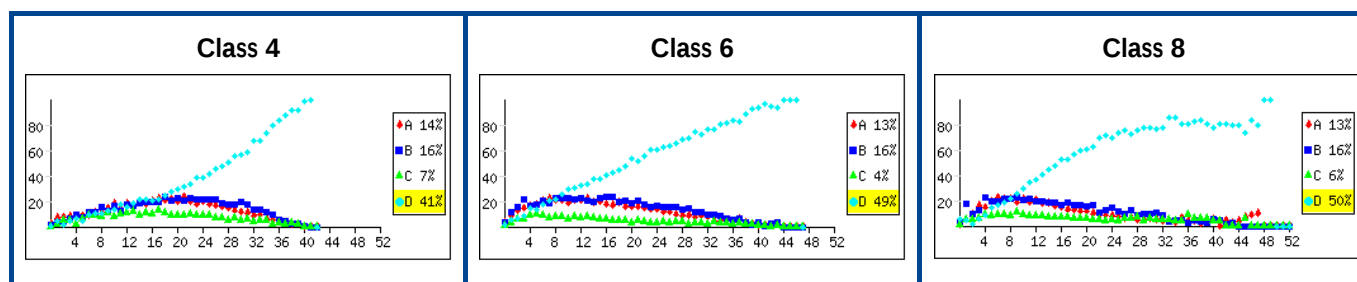
- A. a square
- B. a triangle
- C. a circle
- D. a cube



Sample Question 56: This is a question that involves recognition of basic shapes.

x –classes tested; y – percentage of students who answered correctly.

Analysis of the item response curves of questions reveal that specific patterns of learning exist with groups of students holding on to certain misconceptions at different age levels and ability levels. These can also be seen from the fact that the number of students choosing different options is fairly consistent and are mostly found to vary by 3-4% in different classes.



Graph 2.7: The Item response curves for sample question 54

Graph shown above: These are the item response curves for the sample question 54 that checks for the understanding of the concept of 'half'. The graphs show the percentage of students choosing different answer options and their total scores in the paper. The X axis is the total score in the paper and the Y axis is the percentage of students. The legend shows the total percentage of students who chose each option.

2.4 Free Response Items that required Writing

Question items that required the students to respond in writing were often found to be the most difficult items in the language papers. Further, the scale anchoring analysis which identifies what students at different ability levels know and are able to do identified that the key differentiating skill of students who are high performers in a class are those who are able to handle the questions that require writing.

Based on this, and in order to find out the actual errors and limitations the students face while responding to writing items, a qualitative analysis of answer papers was carried out in Hindi and a team of language experts went through these responses. 30 answer papers for analysis were chosen randomly from each of the 4 districts of 3 Hindi speaking states for classes 4, 6 and 8. A total of 90 answer scripts were analysed in this manner.

2.4.1 Types of questions that required writing

Questions on writing ranged from checking for correct spellings, reorganizing a jumbled sentence or sentences, writing a single to a few meaningful sentences for a picture or a topic, punctuating a sentence, completing a cloze passage where words had to be filled in, completing a dialogue between a shopkeeper and a girl who goes there to buy pencils, and completing a miniature short story by adding 2 sentences from the word clues given.

2.4.2 Key Insights

1. Students are learning language less as a means of natural communication that is used in real life and probably more as a subject. Their writing does not express thoughts beyond the most basic, tried and tested formulaic sentences they probably trained for while in their lower classes.

In the writing tasks, in general, students display a tendency to stick to responses that they have been trained or conditioned to give. When formulating answers, they tend to repeat part of the question stem and not use sentences which are more natural. This comes out even more clearly when we look at responses that required completing a dialogue where half the conversation is already provided. Some of the students have answered using response appropriate to the written form rather than using responses appropriate to a conversation. For example, for the question, "What is the price of this pencil?" the response is, "The price of this pencil is Rs. 4." Rather than only "Rs. 4", as is natural in spoken language.

Class 6, 8 Language

लीला: क्या आपकी दुकान पर पेन्सिलें मिलेंगी?

दुकानदार: जी हाँ मैरी दुकान पर पेन्सिल मिलती है।

लीला: मुझे एक पेन्सिल चाहिए।

दुकानदार: जी दुकानदार ने एक पेन्सिल दी।

लीला: इस पेन्सिल का मूल्य कितना है?

दुकानदार: इस पेन्सिल का मूल्य चार रुपये है।

लीला: यह लीजिए 4 रुपये।

दुकानदार: दुकानदार ने चार रुपये लिये।

Leela: *kya aapkii dukaan par pencillen millengii?* (Do you have pencils available in your shop?)

Shopkeeper: *jii haan merii dukaan par pencillen millthii hai.* (Yes, ma'am. I have pencils in my shop)

Leela: *mujhe ek pencil chahiye.* (I want one pencil.)

Shopkeeper: *dukaandaar ne ek pencil dii* (The shopkeeper gave her a pencil)

Leela: *is pencil kaa muulya kitnaa hai?* (How much does this pencil cost?)

Shopkeeper: *is pencil kaa muulya chaar rupya hai.* (This pencil costs four rupees.)

Sample E: This question asks for the completion of a dialogue between a girl who wants to buy pencils and the shopkeeper.

This is also evident in questions which provide adequate scope to students to express their thoughts, for instance, when writing on familiar topics like "Myself" or "My Favourite Game". Students often do not explore or express thoughts beyond the most basic, tried and tested formulaic sentences they probably trained for in their lower classes. For example, a short paragraph of 3 sentences on "Myself" by class 6 students often are sentences that contain information about their names, age, the class they study in, etc. as in "My name is X. My father's name is Y. I live in Z" etc. At class 8 level, a similar question on "My Favourite Game" often brings out responses like "X is my favourite game. I love playing X. I play X everyday."

Class 6 Language

16 अपने बारे में 3 वाक्य लिखो।

मेरा नाम सुरज कुमार है।
मेरे पिताजी का नाम श्री रामबाबू सिंह है।
मैं चन्दा गाँव में रहता हूँ।

Question: *apne bare men tiin vaakya likho* (Write three sentences about yourself).

mera naam suraj kumar he. (My name is Suraj Kumar.) mera pitaajii ka naam shri rambabu singh he. (My father's name is Rambabu Singh.) mein chandaa gaanv me rahta huun. (I live in Chanda village.)

Sample F: This question checks for the ability to write 3 meaningful and grammatically correct sentences about a given context.

Class 8 Language

17 अपने प्रिय खेल के बारे में 3 वाक्य लिखो।

मेरे प्रिय खेल खो है।
मे प्रति दिन खो खेलती हूँ।
मुझे खो बहुत अच्छी लगती है।

Question: *apne priya khel ke bare men tiin vaakya likho* (Write three sentences about your favourite game).

mera priya khel kho he (My favourite game is Kho.) mein prati din kho khelti huun (I play Kho everyday.) mujhe kho bahut achchhi lagti he (I like Kho very much.)

Sample G: This question checks for ability to write 3 meaningful and grammatically correct sentence about a given context.

II. Students face difficulty in expressing their thoughts in their own words.

Most students find it difficult to perform well on tasks where they are given the freedom to express their thoughts using their own words and expressions. They require a lot of guidance in terms of context and sentence structure to do well in a writing task. The words and expressions used are very simple and common. This is shown using the two samples below.

41 नीचे दिए गए शब्दों की सहायता से अपने 2 वाक्य बनाकर इस कहानी को पूरा करें।

भौंका / बड़ा / कुत्ता / साँप / दूर / भगाया

कड़ी धूप थी। एक आदमी पेड़ की छाँव में बैठा किताब पढ़ रहा था। उसका कुत्ता, उसके पैरों के पास बैठा सो रहा था। अचानक

कड़ी धूप थी। एक आदमी पेड़ की छाँव में बैठा किताब पढ़ रहा था। उसका कुत्ता, उसके पैरों के पास बैठा सो रहा था। अचानक
 उसने देखा कि एक साँप उसके कुत्ते के पास बैठा है।
 उसने दूर से आकर कुत्ते को भौंका और उसे दूर भगाया।
 कुत्ता भौंका और दूर भागा।

Question: niche diye gaye shabdon kii sahaayataa se apne 2 waakya banaakar is kahaani ko puraa Karen

bhaunka/badaa/kutta/saanp/duur/bhagaayaa

kadii dhuup thii. Ek aadmi ped kii chhaanv men baitha kitaab padh rahaa thaa. uskaa kuttaa, uske pairon ke paas baitha so rahaa thaa. achaanak

kadii dhuup thii. Ek aadmi ped kii chhaanv men baitha kitaab padh rahaa thaa. uske kuttaa pairon ke paas baitha kuttaa baitha so rahaa thaa. achaanak use ek saanp aakar usne dur bhagaaya kuttaa bhaunkaa badaa dur bhagaakar kuttaa achaanak so gayaa thaa

Question: Complete the story by adding 2 sentences on your own using the hints given.

bark, big, dog, snake ,away, chased

It was a sunny day. A man was sitting under the shade of a tree reading a book. His dog was sleeping near his feet. Suddenly....

It was a sunny day. A man was sitting under the shade of a tree reading a book. His dog was sleeping near his feet. Suddenly, a snake came there (and)... he chased it away
 dog barked chased it far dog had slept suddenly.

Sample H: This question requires students to understand the given leads for the story and complete it.

41 नीचे दिए गए शब्दों की सहायता से अपने 2 वाक्य बनाकर इस कहानी को पूरा करें।

भौंका / बड़ा / कुत्ता / साँप / दूर / भगाया

कड़ी धूप थी। एक आदमी पेड़ की छाँव में बैठा किताब पढ़ रहा था। उसका कुत्ता, उसके पैरों के पास बैठा सो रहा था। अचानक

वहाँ एक बड़ा साँप आया और कुत्ते के गरिष्ठ पर चढ़ने लगा। कुत्ता भौंका और साँप दूर भागा।
 साँप ने कुत्ते को दूर भगा दिया।

Question: niche diye gaye shabdon kii sahaayataa se apne 2 waakya banaakar is kahaani ko puraa Karen

bhaunka/badaa/kutta/saanp/duur/ bhagaayaa

kadii dhuup thii. Ek aadmi ped kii chhaanv men baitha kitaab padh rahaa thaa. uskaa kuttaa, uske pairon ke paas baitha so rahaa thaa. achaanak ...

wahaan ek badaa saanp aayaa aur kutte ke sharer par chadhne lagaa tab kutta jagaa aur thoda aur bhaunkne lagaa tab woh woh aadmii us saanp ko bhagaaya diya

Question: Complete the story by adding 2 sentences on your own using the hints given.

bark, big, dog, snake ,away, chased

It was a sunny day. A man was sitting under the shade of a tree reading a book. His dog was sleeping near his feet. Suddenly....

A big snake came there and started climbing onto the dog's body then the dog woke up and ran a little and started barking then that man drove that snake away.

Sample I: This question requires students to understand the given leads for the story and complete it

III. Punctuation marks are conspicuous by their absence in the writing of most students.

Students are not using even basic punctuation marks like commas and full stops. This results in long sentences that are difficult for a reader to follow (as can be seen in sample I above). Similarly, the responses on questions that required students to punctuate individual sentences also revealed that students do not understand and follow the rules of punctuation.

Class 6
Language

34 कहानी के अंत में किन दो तरीकों से शेर ने खरगोश को अच्छा महसूस कराने की कोशिश की?

शेर ने खरगोश को डराने के लिए एक गिराया और भुकाम

Question: *kahaani ke ant men kin do tariikon se sher ne khargosh ko achchha mahsuus karaane kii koshish kii* (Write two ways in which the lion tried to make the hare feel better at the end of the story.)

sher ne khargosh ko daraane ke liye giiraya aur bhukamp (To scare the rabbit the lion dropped and earthquake.)

Sample J: This question requires students to understand the passage and answer in their own words. The student has not comprehended the passage well and is also not using any rules of punctuation in the response.

IV. Students commit a number of errors in spelling and grammar.

Table 2.3

Snapshot of errors in spelling and grammar in Hindi	
Common errors in spelling mistakes	Sentence structure errors
The wrong maatra is used or it is displaced	Incomplete sentences
E.g. 'geer' instead of 'gir' 'kii' instead of 'ki' 'gaarii' instead of 'gir'	<i>kumar ke maa ne bole, kumar thum neeche uthar jao thum geer</i> Kumar's mother said Kumar you get down you fall
Maatra is missing or added	Disjoint sentences
E.g. 'jaayega' instead of 'jaayeaga' 'utaar' instead of 'utar' 'aadam' instead of 'aadmii' 'taab' instead of 'tab'	<i>uthar peid par maths chad, necha uthar jaa</i> get down don't climb on the tree, get down
'r' and 'd' used for the other	Run on sentences
E.g. 'char' instead of 'chad' 'pad' instead of 'par'	<i>wahaan ek badaa saanp aayaa aur kutte kii shariir par chadne lagaa tab kuttaa jagaa aur ...</i> A snake came there and started climbing onto the dog's body then the dog woke up and ...

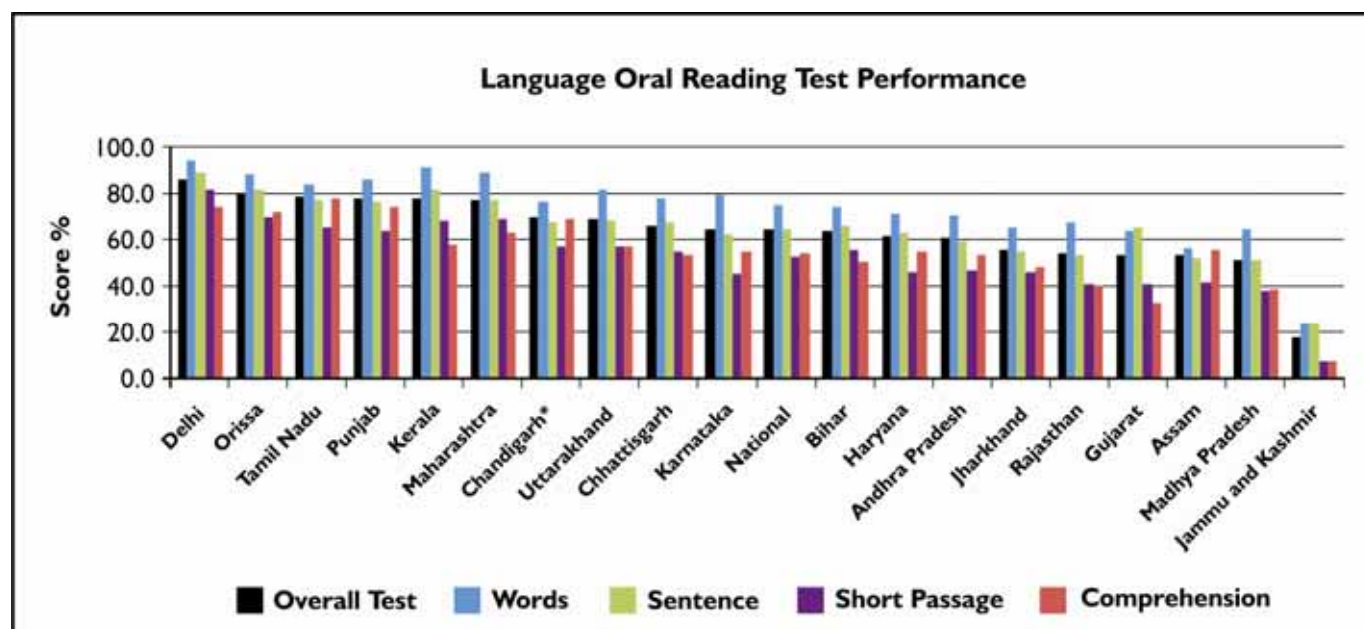
2.4.3 Recommendations

- Students may need to be made to understand vowel signs (maatras) through various activities right from Class 2 onwards; these could involve simple transcription exercises as well as correct pronunciation so it is completely internalized early in the child's educational programme.
- It is important that language learners get to express themselves. A correct answer provided by the teacher that the students learn by rote does not help. They must be encouraged to write new, different sentences using their own words and expressions.
- Students should be given a lot of practise with writing exercises. They can be provided with various stimulus including visuals. These can consist of pictures related to familiar situations. This is essential to the development of a well rounded language user who is able to communicate clearly and fluently.
- These along with constant and positive feedback are important to encourage the learner to learn effectively.

2.5 Language Oral Reading Test

The Language Oral Reading Test was done with a fewer sub sample of students on a one-on-one basis at class 4 level in all the schools tested. About 8 students in each class tested were administered the test. The students were selected in multiples of 3 or 5 by skip counting from the attendance list. Each student was administered the individual oral test alone. The test had a total of 8 questions. The first 3 questions had 9 words each of carefully calibrated difficulty based on word length, familiarity, graphemes (eg, conjunct letters) etc. The 4th and 5th questions had 3 sentences each. The sentences in question 3 were short and simple compared to the sentences in question 4 that were of medium length and had more difficult graphemes in it. The students were asked to read aloud the words and sentences that were pointed out to them by the evaluator. The students also had to read a short passage in question 6. Questions 7 and 8 were comprehension questions based on information stated explicitly and implicitly in the passage respectively. This was done to gain insights into whether the students could decode and comprehend basic information from what they read.

The performance showed that about 60 or more percent of students were able to do all the questions in the test in Delhi, Orissa, Tamil Nadu and Punjab. In all the states tested fewer students comprehended what they were reading. For e.g., in Gujarat while more than 87.1% of students could read a simple word, only 40.3% could read a short passage and only 22.6% could comprehend the information implicit in the passage they read. The scores got by each state in each of the 8 questions are given in Appendix J.



Graph 2.8: The performance of students in Language Oral Reading Test

2.6 Background Factors

This section describes the background questionnaires that were used to collect information related to the student, head-teacher, teacher and school. In particular, it will describe the variables on which the information was collected and their analysis to yield insights on relationships, if any, between these variables and student performance. For purposes of this report, the analysis included students from classes 4, 6 and 8 but for language and maths separately.

2.6.1 Instruments

The instruments are included in Appendix U and V. Table 2.4 below provides a snapshot of the variables on which information were collected.

Table 2.4

Item Context	Variable Description	Questionnaire	Item No.
School Characteristics	Type of school	School	Head
	Multi-grade classroom	School	Head
	Infrastructure	School	101-111

	Material and equipment	School	201-214
	Hours per year	Documents	----
	Mid-day meals program	School	112
Head-teacher background	Gender	Head-teacher	402
	Age	Head-teacher	401
	Academic qualification	Head-teacher	403
	Teacher training qualification	Head-teacher	404
	Years of experience	Head-teacher	406-408
	Type of service	Head-teacher	409
Teacher Background	Gender	Teacher	503
	Age	Teacher	502
	Academic qualification	Teacher	506
	Teacher training qualification	Teacher	507-508
	Years of teaching experience	Teacher	509
	Type of service	Teacher	504
	Perception on discipline	Teacher	517-522
Student Background	Gender	Student	Head
	Age	Student	1
	Socioeconomic background	Student	2,3,4
	Parental occupation	Student	5
	Tuitions	Student	8
Student Perception	Perception about school	Student	6
	Liking for the subject	Student	7
	Use of school library	Student	9
	Student's reading habits	Student	10
	Perception about their own academic performance	Student	11

2.6.2 Analysis Methods

Each student was linked to a teacher and school data based on the combination of school identification and subject and grade taught by the teacher. Thus for comparisons, means were calculated for each category and then t-tests and ANOVA test were conducted to analyze significance of differences.

For regression analysis a sub-sample of the teacher data was used for which student test scores and background data were also available. Since all the analyses were conducted separately for the mathematics and language tests, two separate datasets were created for language and mathematics.

For the student test-scores on mathematics and language, two separate linear regression models were estimated controlling for the student, school, head-teacher and teacher level variables. So the multiple linear regression equation was,

$$Y_{ij} = \alpha_0 + \alpha_1 SchoolVar_{ij} + \alpha_2 HeadteacherVar_{ij} + \alpha_3 TeacherVar_{ij} + \alpha_4 StudentVar_{ij} + \alpha_5 StudentPer_{ij} + e_{ij}$$

Where:

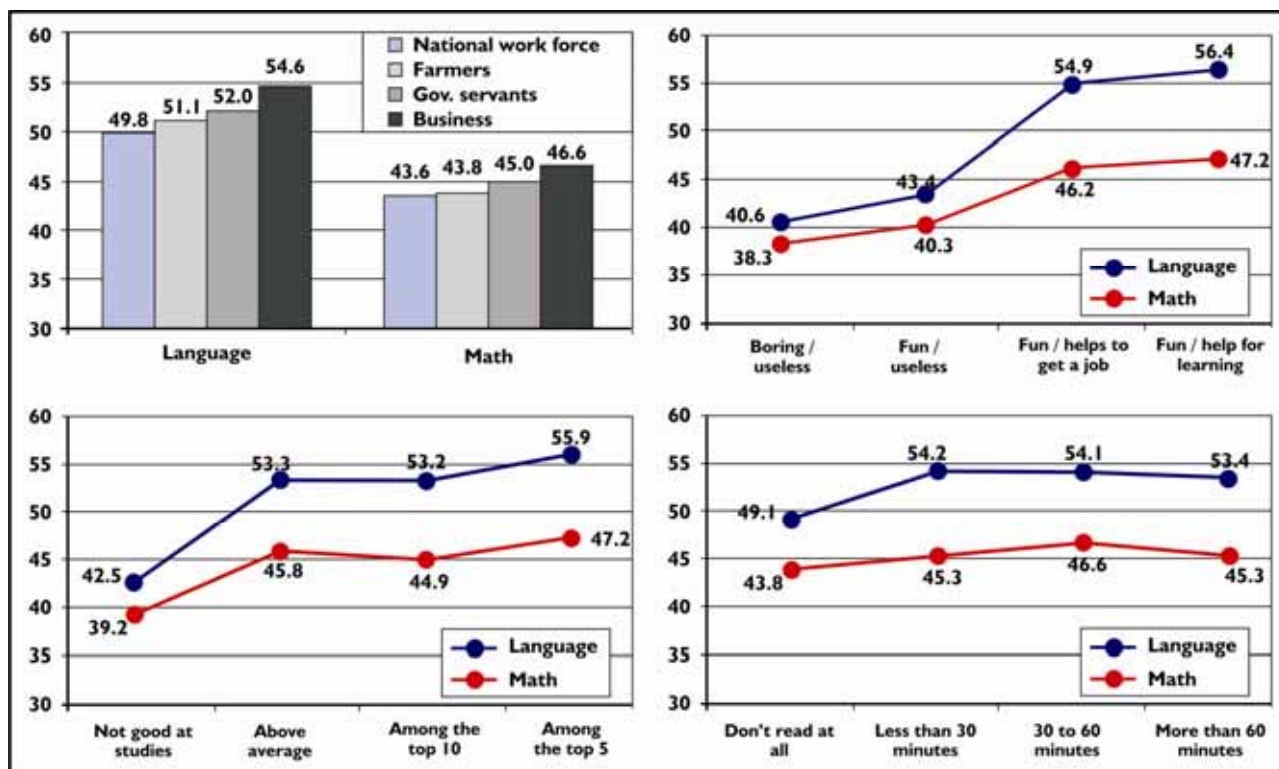
Y_{ij} is the dependent variable, language or maths score. School characteristics ('SchoolVar') included infrastructure, learning resources, presence of mid-day meals, type of school and multi-grade classes. Head-teacher variables ('HeadteacherVar') comprised age, sex, type of service, educational qualification and years of experience. Teacher background ('TeacherVar') included age, sex, type of service, educational qualification, years of experience and beliefs about discipline. Student background ('StudentVar') incorporated sex, age, socioeconomic background, occupation of parents and presence of tuition for subjects. Student perception and habits ('StudentPer') included preferred subject, perception about school, perception about their own performance, daily time spent on readings and use of library school: "e" represented the error term and α the coefficient vectors of the model. Finally, these regression equations were also analyzed separately for different classes to verify certain findings by subgroups.

2.6.3 Analysis of the student background and their perceptions

In addition to gender and age, ten questions about student background and their perceptions were included in the student questionnaire administered at the end of the tests. Further analysis was carried out on the dataset of language and mathematics students separately.

1. Gender: While there are no differences between the average scores of boys (51.9) and girls (51.9) for language; the average score of boys is significantly higher (46.0) than the average score of girls (42.7) for maths.

2. Age: There is a negative correlation between age and maths scores. It means that the higher the age, the lower the score. Further detailed analysis by class indicates that this is due to the presence of students of higher ages than would be expected as appropriate to the class. As could be expected, students who do not progress normally as per age to higher classes might be low performing students.



Graph 2.9: Means of students' scores in language and maths crossed with variables in Student Questionnaire

3. Parental occupation: For maths and language respectively, children whose parents were involved in business (54.6 and 46.6) performed better than children of government servants (52.0 and 45.0); and government servants' children were seen to perform statistically better than the children of farmers (51.1 and 43.8). Particularly for language, children of farmers scored higher than children of parents from the national work force (49.8). In maths, there were no differences between these two groups in statistical terms.

4. Student's opinion about learning: As can be seen in the graph above, students who think school is not useful obtain lower scores in both subjects (40.6 and 38.3). In contrast, students who perceive school as fun and useful reached higher scores. The percentage of students who consider coming to school as useless is around 16%.

5. Student's perception of their own performance: Students who see themselves as being bad at studying show the lowest average score for language (42.5) and maths (39.2). In contrast, students who perceive themselves to be among the top 5 in the class achieved the highest average scores (55.9 and 47.2).

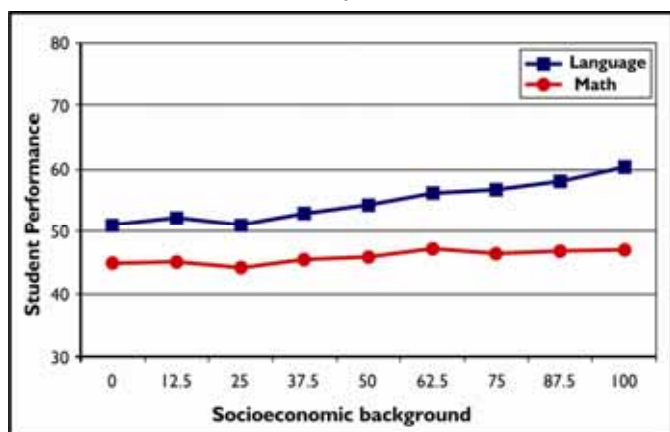
6. Student's liking for the subject: Students who did not like any subject reported significant lower scores in both language and maths. At the same time, students who liked maths more reported significant higher scores in that subject. Students who liked language more achieved higher scores in that subject, but their score was not significantly different from students liking maths.

7. Students' reading habits: Students who reported not to have the habit of reading material other than textbooks achieved significantly the lowest scores for both language (49.1) and maths (43.8). Additionally, reading more than 60 minutes does not show additional improvements in the student performance.

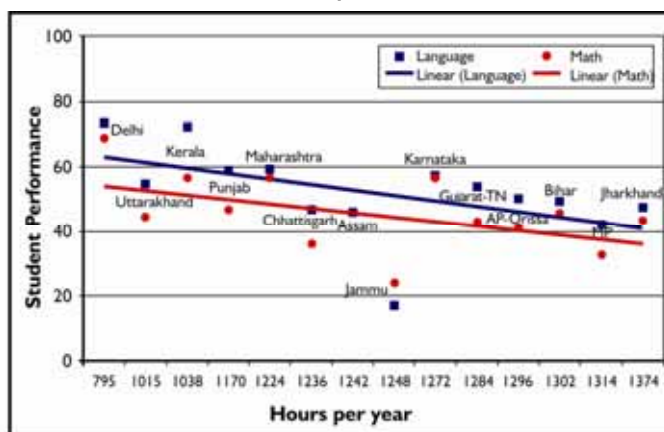
8. Students taking tuitions: Students who did not take tuitions reported significantly better performance (56.4) than students taking tuitions (49.3 in maths, 49.1 in language and 54.3 in both subjects). A hypothesis to explain this result is that students take tuitions because they needed complementary support due to deficient performance. The percentage of students who did not take tuitions was 30.7 while 17.6% of students affirmed that they take tuitions for maths; 14.5% for language and 37.2% take for both subjects.

9. Socioeconomic background: There is a modest relationship between socioeconomic background¹⁰ and student performance for language. The higher the socioeconomic background, the higher the language scores as can be seen in the graph below. The relationship between the maths score and socioeconomic background was weak.

Graph 2.10



Graph 2.11

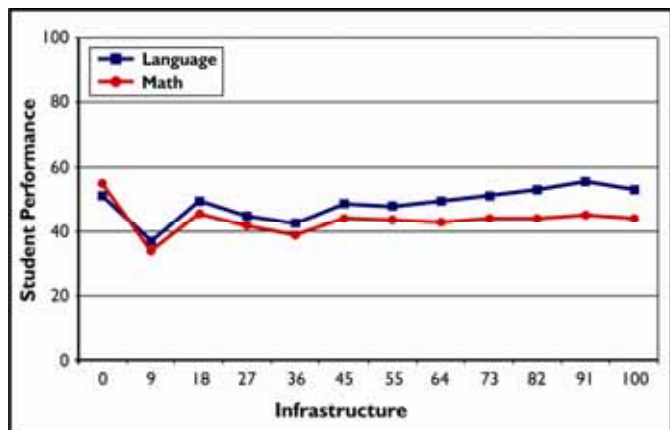


10. Number of academic hours per year: *Increasing the number of academic hours per year does not guarantee better student performance.* As can be seen from graph 2.11, the relationship between the number of working hours and student performance is indirect. It means, for instance, that schools that invested around 1300 hours per year developed lower average score for language and maths than schools that invest around 800 hours.

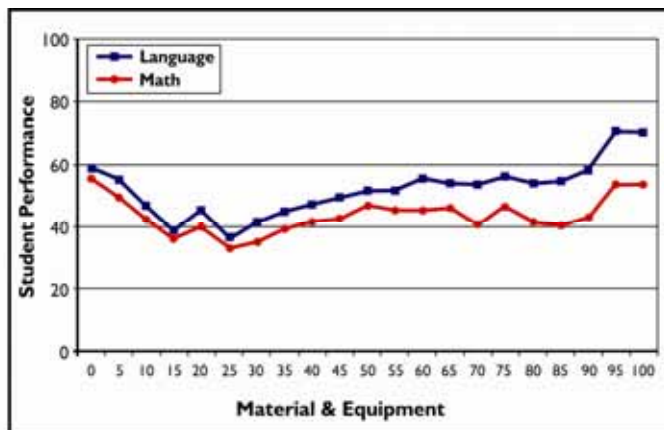
2.6.4 Analysis of the school background

How much do the infrastructure and the learning resources matter for the student performance? To answer this enquiry, 12 dichotomous questions regarding infrastructure and 14 questions on materials and equipment were collected. Based on simple additive scores, two scales – one for infrastructure and another for material and equipment – were computed to establish the relationship between infrastructure, learning resources and student performance. Average scores for language and maths by each question are shown in Appendix I.

Graph 2.12



Graph 2.13



¹⁰Socioeconomic background for this report is measured by an additive scale of three variables based on possession of belongings such as cycle, scooter, radio, television and stove.

11. Infrastructure: *Although the relationship is positive, the relationship between infrastructure and student score is weak.* Correlations indicate that schools that have basic infrastructure and services such water, electricity and toilets do not necessarily ensure better performance.

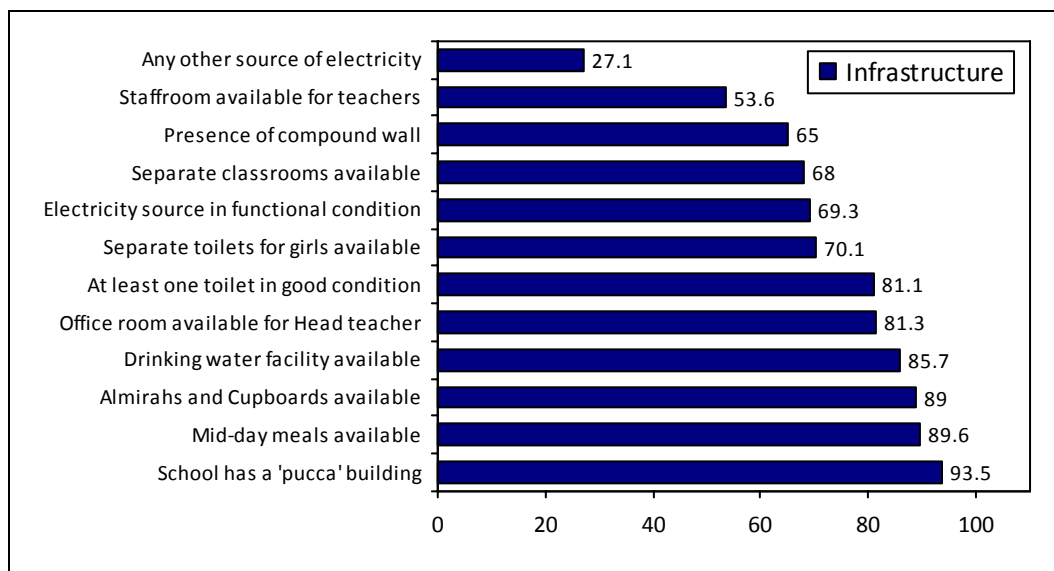
12. Learning resources: *The data reveals that the presence of material and equipment has a modest relevance for increasing the achievement of students in language. In maths, the relationship between learning resources and learning scores is low.* For language, the presence – in that order – of books, functional computers, available computers for students and blackboards would contribute towards obtaining better scores. For maths, blackboards are the most important facility to improve student learning.

13. Mid-day meals: For language and maths, students whose schools provided mid-day meals performed significantly better (51.8 and 44.8) than those whose schools did not provide daily meals (49.1 and 39.4). Note that 89% of the students received daily meals.

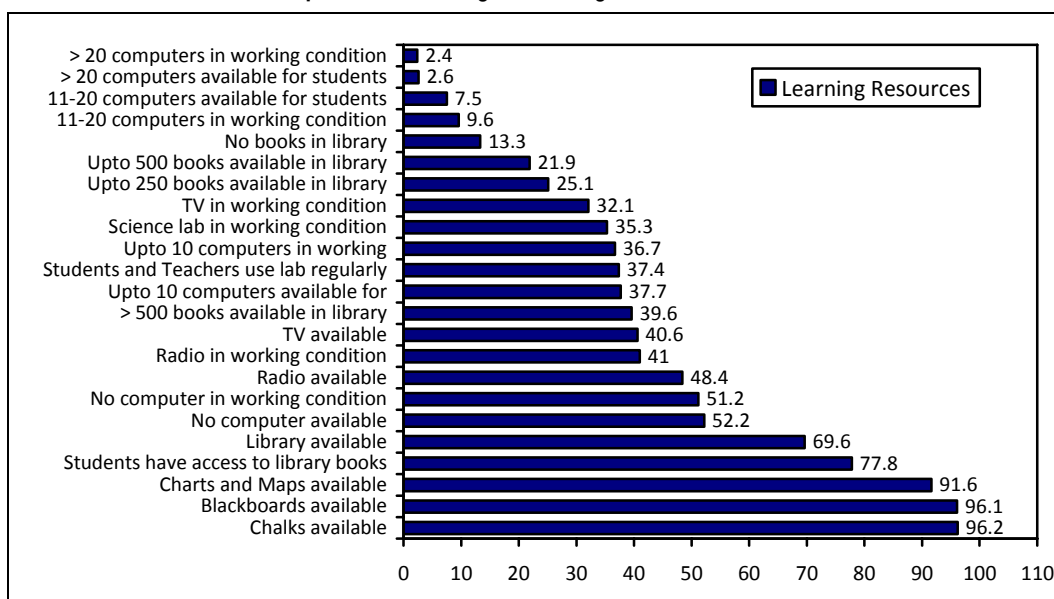
14. Type of school: Although the difference is not very large, students from girls' schools performed significantly lower than students from boys' and co-educational schools for both subjects.

15. Multigrade classes: Students from multigrade schools achieved higher average scores (50.9 and 44.0) than students from monograde schools in language and maths respectively (49.4 and 42.0).

Graph 2.14: Percentage of Infrastructure Characteristics of Schools



Graph 2.15: Percentage of learning resources in schools



2.6.5 Analysis of the head-teacher background

Information regarding the head-teachers was analysed to understand their role in student performance.

16. Gender: Analysis shows that students in schools managed by a woman achieved significantly higher average language and maths scores (55.3 and 47.2) than students in schools managed by a man (49.8 and 42.8).

17. Type of Service: Students from schools that had head-teachers appointed on a contract-basis perform statistically better than students from schools with regular head-teachers in language, while there are no such differences in maths.

18. Academic Qualification: When student scores are compared by highest degree of the head-teacher, it is seen that students of head-teachers with higher academic qualification perform significantly lower. Students whose head-teacher had a master's degree scored (40.4) lower than students whose head-teacher had a bachelor's degree (43.7), who in turn performed lower than students whose head-teachers did not have any degree (53.9).

19. Teacher Training Qualification: Students whose head-teachers did not have teacher training qualifications achieved significantly lower scores in language (46.9) than students whose head-teachers have a bachelor's degree (51.0), diploma (49.9) or master's degree in education (49.5). For maths, students whose head-teachers had a master's degree in education achieved the lowest performance (40.6).

20. Experience and Age: There are no significant relationships between experience as head-master or age of the headmaster and the performance of the students in language and maths.

2.6.6 Analysis of the teacher background

21. Gender: Students of female teachers performed significantly better in language (53.3) than students of male teachers (48.7).

22. Type of Service: The data reveals that contracted teachers' students perform statistically better (51.2) than regular teachers' (45.9)

23. Academic Qualification: Students of teachers without a degree score higher than students of teachers with bachelor's or master's degrees for language and maths. In language, students of teachers without a degree, with a bachelor's degree and with a master's degree achieved 52.4, 50.4 and 49.4 respectively. Similarly for maths, students of teachers without a degree, with a bachelor's degree and with a master's degree achieved 51.9, 42.5 and 40.5 respectively.

24. Teacher Training Qualification: Students of teachers without teacher training achieved 45.2 in language while students of teachers with diploma in teacher training, bachelor's in Education and master's in Education scored 50.2, 49.2 and 49.0 respectively. For maths, students of teachers who do not have teacher training achieved 37.0 while students of teachers with a diploma in teacher training, bachelor's in Education and master's in Education scored 46.6, 40.6 and 41.1 respectively.

25. Teacher Experience and Age: There is no correlation between years of teaching experience or age and performance of their students.

2.6.7 Interaction of background variables: Regression analysis

Further investigations were carried out to check if any of the student, school, head-teacher or teacher factors could be used to predict the performance of students. This was done by carrying out multiple linear regression analysis. Methodological aspects and tables are detailed in Appendix I.

As could be seen from the regression analysis, the most relevant factors to influence student learning were:

1. Students perceiving themselves as good is among the top 5 influences on student achievement by 9.1 and 5.2 points for language and maths respectively, compared to students who feel that they are not good at studies.

2. Perception of the school by student – compared to students who perceive the school boring and not useful, students who perceive it as fun and useful to learn new things or to find a job in the future influences student achievement by 8.8 and 9.2 respectively in language and 5.8 and 6.3 in maths.
3. Studying at a girl-only school negatively influences student achievement by 10.8 (language) and 9.3 (maths) points if compared to co-educational schools.
4. Teachers with a master's degree negatively influence student achievement by 8.5 points for language and 3.2 points for maths when compared to teachers without a degree.
5. Teachers with a diploma in teacher training positively influence maths learning by 7.0 points while those with bachelor's in Education and master's in Education influence by 4.2 and 6.2 points respectively.
6. Head-teachers with an academic degree negatively influence students learning in language and maths when compared to head-teachers without a degree.

Table 2.5

Independent Variable	Language	Maths
1. School characteristics		
Infrastructure	-0.04*	0.04*
Material and equipment	0.24***	0.06***
Mid-day meals	-2.64*	-0.64
Academic hours per year	-0.05***	-0.05***
Type of school ¹		
Boys school	0.25	-0.25
Girls school	-10.77***	-9.34***
Monograde school ²	-0.72	-0.02
2. Head-Teacher Background		
Age of teacher	-0.18***	-0.17***
Headmistress ³	-0.26	-0.11
Type of service ⁴		
Permanent	-2.06*	-7.25***
Educational qualification ⁵		
Bachelor	-3.62**	-10.44***
Master	-3.56**	-12.54***
Teacher training qualification ⁶		
Bachelor in Education	0.33	-3.78***
Diploma in Education	-3.16**	-4.61***
Master in Education	-2.93*	-4.28***
Training in the last year	0.68	0.64
Years of Experience as a teacher	0.18***	0.20***
Years of Experience as head-teacher	0.03	-0.37***
Years of Experience as head-teacher at this school	0.15	-0.08
3. Teacher Background		
Age of teacher	0.00	0.09
Female teachers ⁷	2.94***	2.53***
Type of service ⁴		
Regular	-0.52	2.12**
Educational qualification ⁵		
Bachelor	-5.93***	-2.84**
Master	-8.15***	-3.33***
Teacher training qualification ⁶		
Bachelor's in Education	-1.50	4.22***
Diploma in Education	-1.11	6.95***
Master's in Education	-2.39	6.18***
Training in the last year	2.82***	0.62

Years of experience	0.03	-0.19***
Handling multigrade classroom	-1.10*	3.26***
Perception about discipline	0.60**	-0.22
4. Student Background		
Male student ¹⁵	-0.88	3.03***
Age	-0.69***	-2.33***
Socioeconomic background	0.05***	0.05***
Parents' occupation ⁸		
Farmers	1.07*	0.17
Government servants	-0.48	1.31
Business people or private workers	2.32*	2.21**
Tuition for subject ⁹		
Maths only	-4.36***	-5.77***
English only	-5.78***	-4.97***
Maths and English	-2.57***	-2.86***
5. Student perceptions and habits		
Perception about school ¹⁰		
Fun and helps to learn new things	8.88***	5.80***
Fun and help to find a job when growing up	9.23***	6.34***
Fun, but not useful	-1.15	-0.22
Preferred subject ¹¹		
English	2.79**	0.72
Maths	2.61*	2.63**
Another subject	2.08*	-1.17
Spend time daily on reading material other than textbooks ¹²		
more than 1 hour	0.44	0.40
30 minutes to 1 hour	1.86*	2.53***
less than 30 minutes	2.36**	2.12**
Use of school library ¹³		
Library doesn't give books	-3.55***	-1.67*
Don't take books	-0.82	0.73
Take books	3.76***	2.40***
Perception about their performance ¹⁴		
Among the top 5 in the class	9.13***	5.20***
Among the top 10	6.38***	4.53***
Above average	7.34***	4.29***
TOTAL	7290	7715
R ²	29.2	30.1

Note:

- *** = $p \leq .001$, ** = $p \leq .005$, * = $p \leq .01$, + = $p \leq .05$
- References categories:** 1 = Coeducational school, 2 = Multigrade schools, 3 = Headmaster, 4 = Contract, 5 = No degree, 6 = No teacher training, 7 = Male teachers, 8 = Army, navy or air force, 9 = No tuitions, 10 = Boring and not useful, 11 = Don't like any subject, 12 = Don't read at all, 13 = School without library, 14 = Not good at studies, 15 = Female Students.

2.6.7 Discussion

Providing students with solid knowledge and cognitive skills as well as democratic and moral values are often the goals of public school systems. Particularly in India, the Government of India's flagship programme Sarva Shiksha Abhiyan (SSA) has taken on the challenge of providing universal and quality education at elementary level. Among others, the programme seeks to strengthen existing school infrastructure through provision of additional class rooms, toilets, drinking water, maintenance

grant and school improvement grants, while the capacity of existing teachers is being strengthened by extensive training and grants for developing teaching-learning materials (SSA, 2009). However, conclusive and consistent evidence on factors that determine student outcomes is scarce in developing countries such as India. Any analysis for predicting outcomes based on observed background variables suffer limitations in terms of not including, in the questionnaire, variables that are not logically apparent as ones that could have possible relationship with student learning.

The factors included in this research indicate that the percentage of score variance explained by the regression model was 29.2% for language and 30.1% for maths. As was observed in Bhutan (Educational Initiatives, 2009), student perception about themselves and on the school are good factors to predict the student performance.

However, in terms of policy it is important to understand what is necessary to do at the school level. Among them, teacher and teacher quality has been identified as the most important factors associated with student achievement (UNESCO, 2004). Because teacher quality is difficult to define, teaching qualification is used as a proxy indicator. The McKinsey Report (2008) on the best performing school systems mentions that the quality of an education system cannot exceed the quality of its teachers and the only way to improve outcomes is to enable teachers to become effective instructors.

The comparison of means of student scores indicated that teacher training is important for maths and language. However, after controlling for the effect of other associated factors, regression analysis revealed that the effect of teacher training is specifically relevant to achieve maths learning. Comparison of means of student scores and regression analysis showed that students of teachers with an academic degree achieved lower scores in language and maths. These suggest that the recruitment and the training system for teachers have to be reviewed in order to improve the quality of instruction in these schools.

Research also suggests that quality goals need to be set along with quantity goals. Effectively, India has shown consistent progress towards universal primary education, but the quality of learning is a pending agenda (PRATHAM, 2005, 2006, 2007, 2008). For example, although many states spend more than 1000 hours of schooling per year as per World Bank estimates (The World Bank, 2004), the increase in the number of hours does not guarantee the progress of quality learning. Moreover, this research found that special interest must be given to girl schools whose students reached lower scores than other students.

Findings reveal the importance of the presence of school libraries and its use by the students for improving student performance. Nevertheless, around 30% of the schools affirmed that they did not have a library. Additionally, analysis shows the relevance of reading habits to be inculcated in students. Spending 30 minutes or less each day to read material other than textbooks is enough to achieve an increase in student learning. Research indicates that if students had more than 100 books in their home, their average achievement would be higher than those who had less than 10 books (Mullis, Martin, Kennedy, & Pierre, 2007). Methodologically, further analysis could include multilevel analysis to discriminate residual effects of student, school and state.

Chapter 3. COMPARATIVE FINDINGS



3.1 Performance of Different States

The data generated based on the tests was analysed to get a comparative picture of the learning levels of students across all 18 states and 1 union territory of India. The analysis is a rich source of information about which states are doing better in terms of overall student achievement than others.

3.1.1 Overall Performance

It must be noted that the performance of the states varies across classes and subjects. We shall therefore take only one subject and class into consideration for interpretation here (language Class 6). The table below gives the number of students in each state and the respective averages with the standard deviation. For similar tables for language and maths of all classes tested, refer to Appendix D.

Table 3.1

Class 6 Language								
S. No	States	N	AVG	SD	Standard Error	Significance	Cohen's <i>d</i>	Effect Size - Interpretation
1	Kerala	1536	72.9	18.6	0.475	YES	1.117	Large
2	Maharashtra	3943	60.5	16.7	0.265	YES	0.547	Medium
3	Chandigarh *	760	58.0	13.8	0.500	YES	0.435	Medium
4	Punjab	1351	57.8	16.8	0.457	YES	0.428	Medium
5	Orissa	1281	55.5	22.7	0.634	YES	0.321	Small
6	Karnataka	2047	55.1	18.5	0.408	YES	0.301	Small
7	Tamil Nadu	3203	51.2	19.8	0.350	YES	0.123	-
8	Uttarakhand	1257	50.2	19.0	0.535	YES	0.079	-
9	Haryana	1431	49.4	19.9	0.526	NO	-	-
10	National Average	35945	48.5	21.9	0.115			
11	Gujarat	1641	47.6	20.7	0.511	NO	-	-
12	Bihar	2734	46.8	21.7	0.415	YES	0.077	-
13	Jharkhand	1269	46.0	19.7	0.552	YES	0.114	-
14	Andhra Pradesh	3789	45.7	18.8	0.305	YES	0.127	-
15	Chhattisgarh	1788	42.6	19.4	0.458	YES	0.270	Small
16	Assam	1060	41.5	20.7	0.636	YES	0.319	Small
17	Madhya Pradesh	3294	37.4	21.6	0.376	YES	0.506	Medium
18	Rajasthan	2618	35.9	19.0	0.372	YES	0.576	Medium
19	Jammu and Kashmir	943	14.4	8.4	0.274	YES	1.557	Large

N - Number of students; *AVG* - Average Score; *SD* - Standard deviation; *Statistical Significance* at 95%; *Cohen's d* - Large: >0.8, Medium: >0.5, Small >0.2

* - Union Territory

Based on the table above, we can say that students from, say, Kerala, Maharashtra, Chandigarh and Punjab, on the whole are clearly performing better than the national average in class 6 language. Jammu and Kashmir, Rajasthan and Madhya Pradesh, were among the states that ranked among the bottom in the language class 6.

Tamil Nadu, Uttarakhand, Haryana, Gujarat, Bihar, Jharkhand and Andhra Pradesh are not meaningfully different from national average.

The SD/Standard deviation of Jammu and Kashmir is noticeably very low showing that the students here are performing very similarly. This when read with the low average score shows that the students from this state are displaying uniformly low performance.

3.1.2 Multiple Comparisons of Performance of Different States

The comparative picture of the performance of different states in Class 6 language and the significance of the difference in performance is represented below. Similar comparison for other classes and subjects are available in Appendix E.

Class-6 Language	Kerala	Maharashtra	Chandigarh *	Punjab	Orissa	Karnataka	Tamil Nadu	Uttarakhand	Haryana	Gujarat	Bihar	Jharkhand	Andhra Pradesh	Chhattisgarh	Assam	Madhya Pradesh	Rajasthan	Jammu and Kashmir
Kerala		▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Maharashtra	▼*		▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Chandigarh *	▼*	▼		▲	▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Punjab	▼*	▼*	▼		▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Orissa	▼*	▼*	▼	▼		▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Karnataka	▼*	▼*	▼*	▼*	▼		▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Tamil Nadu	▼*	▼*	▼*	▼*	▼*	▼*		▲	▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Uttarakhand	▼*	▼*	▼*	▼*	▼*	▼*	▼		▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Haryana	▼*	▼*	▼*	▼*	▼*	▼*	▼	▼		▲	▲*	▲*	▲*	▲*	▲*	▲*	▲*	▲*
Gujarat	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼		▲	▲	▲	▲*	▲*	▲*	▲*	▲*
Bihar	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼		▲	▲	▲*	▲*	▲*	▲*	▲*
Jharkhand	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼	▼		▲	▲*	▲*	▲*	▲*	▲*
Andhra Pradesh	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼	▼	▼		▲*	▲*	▲*	▲*	▲*
Chhattisgarh	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*		▲	▲*	▲*	▲*
Assam	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼		▲*	▲*	▲*
Madhya Pradesh	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*		▲	▲*
Rajasthan	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼		▲*
Jammu and Kashmir	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	▼*	

* The mean difference is significant at the .05 level.

Legend:

- ▲* - Average achievement of State in column 1 is significantly higher than comparison State.
- ▼* - Average achievement of State in column 1 is significantly lower than comparison State.
- ▲ - Average achievement of State in column 1 is higher than comparison State, but not significant.
- ▼ - Average achievement of State in column 1 is lower than comparison State, but not significant.

The chart above shows that the average achievement of schools of Kerala is significantly higher than those of all other states, while the average achievement of schools of Jammu and Kashmir is significantly lower than those of all other states. The schools of Gujarat are performing significantly higher than schools of Chhattisgarh, Assam, Madhya Pradesh, Rajasthan and Jammu & Kashmir. Gujarat is performing better but not significantly better than Bihar, Jharkhand and Andhra Pradesh. In Punjab, schools are performing significantly lower than schools of Kerala and Maharashtra. Schools in Punjab are performing

lower, but not significantly so, than schools of Chandigarh. Similar explanations can be derived by reading the table above for other states with the help of the legend given below the table.

3.2 Comparative Performance of Boys and Girls

Over 51,500 girls and 49,000 boys participated in the study and an analysis was done to find out if there was a noticeable difference between the average performance of the boys and girls in the study. The table below indicates the mean difference in the performance according to gender:

Table 3.2

Class	Subject	Number of Girls	Number of Boys	Girls (%) correct	Boys (%) correct	Significance	Cohen's d	Effect Size - Interpretation
4	Language	14859	14719	55.1	54.6	0.053	0.022	-
6	Language	18553	17392	48.6	48.4	0.301	0.009	-
8	Language	18586	17534	53.1	53.8	0.000*	-0.037	-
4	Maths	14827	14686	55.5	57.6	0.000*	-0.096	-
6	Maths	18371	17233	42.4	46.4	0.000*	-0.182	-
8	Maths	18543	17424	33.7	37.1	0.000*	-0.219	Small

N - Number of students; *AVG* – Average Score; *SD* – Standard deviation; *Statistical Significance* at 95%; *Cohen's d* – Large: >0.8, Medium: >0.5, Small >0.2

Statistical analysis indicated that in language, boys are performing significantly better than girls in Class 8 language and maths 4, 6 and 8. However, as the number of students tested is very large and may by itself influence the statistical significance, Cohen's *d* was carried out further to check for the meaningfulness of the differences observed. **It can be seen that in Maths 8, boys are doing significantly better than girls, and the magnitude of the difference is meaningful.** The pictorial representation explains this below.



The comparative performance of boys and girls is similar to many international studies – boys seem to do better in Maths, although the SLS study reveals that it is a meaningful difference/a difference that matters in class 8.

It should be stressed that this is probably due to cultural reasons and the findings do not suggest that boys are inherently more capable of doing Maths. The differences in performance indicate, if anything, the cultural and social differences in the upbringing of boys and girls in terms of encouraging boys to study Maths and discouraging girls from doing so. This disparity can be corrected if the government, teachers and schools take steps to encourage girls who are interested in taking up Maths. Moreover, the society's role in the development of both the genders should be recognised in order to root out its negative influences and encourage its potential positive impacts.

3.3 Comparison Across Urban and Rural India

The table below gives the number of students in each category and the respective averages with the standard deviation for classes 4, 6 and 8 for language and maths.

The school enrolment data collected for districts of Chandigarh (class 4, 6 and 8) and Surendranagar (for class 8) did not distinguish the schools as urban or rural and hence are excluded from the above analysis.

Table 3.3

Subject	Urban			Rural			Significance	Cohen's d	Effect Size - Interpretation
	N	AVG	SD	N	AVG	SD			
L4	7212	58.9	21.3	21635	53.3	23.5	0.000*	0.242	Small
L6	7144	52.6	20.2	28041	47.2	22.3	0.000*	0.247	Small
L8	7983	55.5	17.8	26256	52.5	19.4	0.000*	0.157	
M4	7184	58.4	19.8	21598	55.9	22.6	0.000*	0.114	
M6	7049	47.2	20.8	27795	43.5	22.4	0.000*	0.168	
M8	7959	35.3	14.7	26127	35.4	15.8	0.777	-0.006	

N - Number of students; AVG – Average Score; SD – Standard deviation; Statistical Significance at 95%; Cohen's d –Large: >0.8, Medium:>0.5, Small >0.2

As could be seen from the table above, except for Maths 8, the differences between urban and rural students are statistically significantly in all classes and subjects tested. However, as could be seen from the Cohen's d, the difference is meaningful only in language 4 and 6, with urban students performing better than their rural counterparts.

3.3.1 State-wise Urban and Rural Performance

The table below shows the performance of urban and rural schools in each state for language class 4. In most of the states urban schools are performing better than rural schools except the states of Jharkhand, Chhattisgarh and Maharashtra. In the states of Bihar, Gujarat, Jammu and Kashmir, and Uttarakhand the difference between the performance of urban schools and rural schools is not statistically significant. The analysis for other class and subjects is given in the Appendix F

Table 3.4

S. No	State	Urban			Rural			Significance	Cohen's D	Effect Size Interpretation
		N	AVG	SD	N	AVG	SD			
1	Andhra Pradesh	746	54.7	18.9	1619	50.9	21.2	0.000*	0.184	
2	Assam	91	55.7	21.2	496	45.5	23.5	0.000*	0.436	Small
3	Bihar	580	49.6	25.3	1969	49.5	23.6	0.900		
4	Chhattisgarh	229	46.9	20.8	1175	51.7	20.0	0.001*	-0.238	Small
5	Delhi	823	74.6	15.9	112	66.8	20.2	0.000*	0.470	Small
6	Gujarat	484	50.5	21.8	1206	51.0	23.1	0.649		
7	Haryana	121	61.2	19.9	1064	49.3	25.3	0.000*	0.474	Small
8	Jammu and Kashmir	58	19.4	10.2	396	18.9	13.3	0.728		
9	Jharkhand	117	42.8	20.9	888	48.9	22.4	0.004*	-0.274	Small
10	Karnataka	1053	60.9	17.1	824	55.4	20.8	0.000*	0.289	Small
11	Kerala	164	72.9	17.9	1029	69.5	20.1	0.030*	0.171	
12	Madhya Pradesh	503	54.0	22.3	1555	37.9	26.4	0.000*	0.610	Medium
13	Maharashtra	527	54.0	18.2	3454	64.5	17.3	0.000*	-0.590	Medium
14	Orissa	220	67.1	19.9	555	62.9	23.3	0.013*	0.187	
15	Punjab	147	64.0	18.2	924	56.3	18.6	0.000*	0.410	Small
16	Rajasthan	162	50.4	21.0	1404	39.8	22.9	0.000*	0.461	Small
17	Tamil Nadu	1054	64.6	17.2	2116	59.5	19.3	0.000*	0.271	Small
18	Uttarakhand	133	60.8	21.5	849	57.6	20.7	0.113		

3.4 Comparison With National and International Performance

Among the questions used in the SLS 2009 there were a few carefully selected items from international tests – questions on which the performance of students in different parts of the world was already known. The table below shows the number of such questions in different papers. Questions were taken from tests that have been done on over 10,000 students each in private schools of India, government schools of Bhutan and International students. The international tests from which the

questions were taken are the Trends in International Maths and Science Study (TIMSS) and the Progress in Reading Literacy Study (PIRLS) (please refer Chapter 1 for details.)

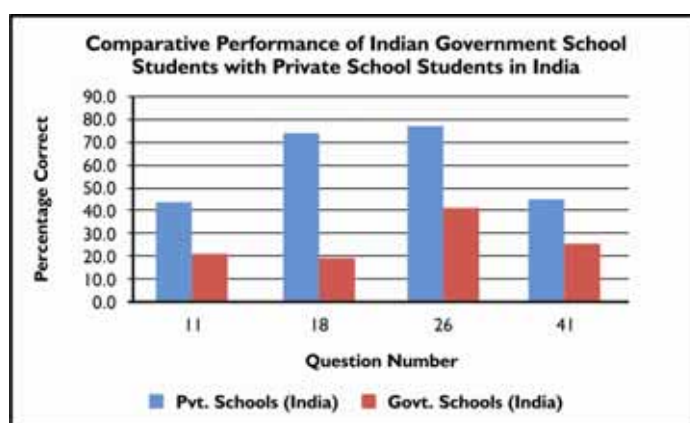
The table 3.5 below gives the questions that were commonly used from the other studies as anchor items

Table 3.5

Anchor Items Chosen From	Language 4	Language 6	Language 8	Maths 4	Maths 6	Maths 8
Benchmarking Studies for Private Schools in India	0	0	0	0	0	4
International Tests (TIMSS & PIRLS)	5	6	6	7	0	7
Benchmarking Studies for Government Schools in Bhutan	5	8	0	0	0	4

3.4.1 Benchmarking Performance with Private Schools in India

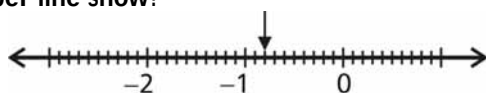
The performance of government school students in the study was compared with private school students in class 8 maths. A total of 4 questions were used for the comparison. The private school students outperformed the students of government schools in each of the 4 questions. The gap is wide between the 2 school systems in all the questions shown, with the maximum gap observed in question 18, where Indian private school students have performed at 73.5% and Indian government school students have performed at 19% only.



Graph 3.1: The performance of students on common questions

Class 8 Maths

Which rational number does the pointer on the number line show?



- A. -1.2
- B. -0.4
- C. 1.2
- D. -0.8

Correct Answer: D
Performance of Class 8 Students

Group	% correct
Pvt. Schools (India)	73.5
Govt. School (India)	19.0

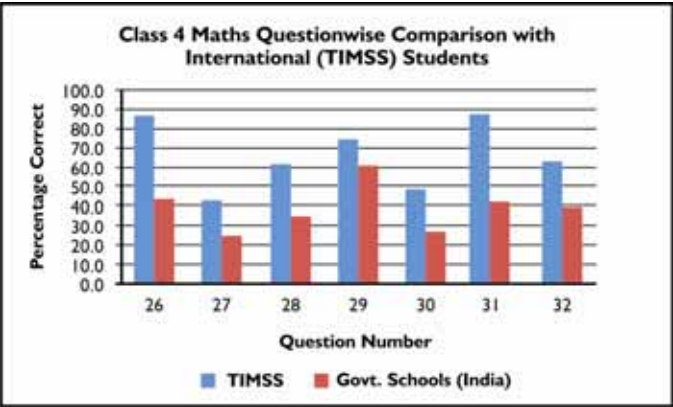
Sample Question 57 Only about **19.0%** of the students chose the correct option D (-0.8). The most common wrong answers selected are option A (40.8%) and option C (22.9%). This shows that students are finding it difficult to understand the sequencing of numbers on a number line. However a different trend is observed in private schools **73.5%** of students are able to answer this question correctly.

3.4.2 Benchmarking Performance with Students in Other Countries

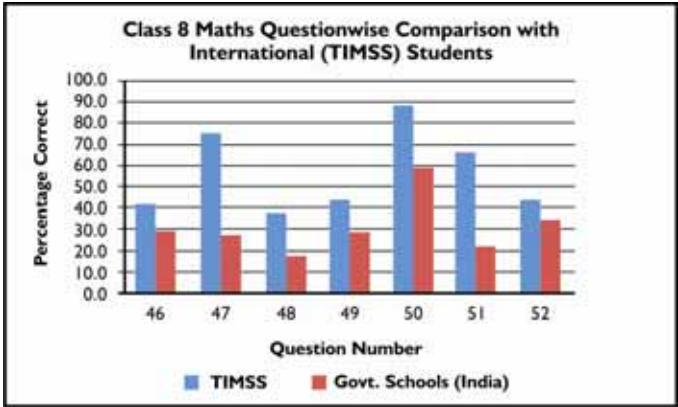
The common questions used to benchmark the performance with international students were taken from the TIMSS – an international research project of high repute. These tests are conducted every 4 years, and some of the questions are publicly released along with their performance data. About 11,000 students from over 40 countries have taken these tests. These

questions were mostly used with no modification, except that 1. names were changed to Indian names and 2. from some questions that had 5 options, the least probable option was removed to make them fit the format of the questions of this study. Similarly, the PIRLS is an international test on language and literacy skills. TIMSS items were used in grades 4 and 8, and PIRLS items in grade 4. PIRLS class 4 items were also used in Class 6 and 8 papers to see how students in higher classes in India are able to perform in comparison to international class 4 students.

In class 4 and 8 maths papers that were benchmarked, the performance of Indian students is below the TIMSS students, in ALL the questions commonly used from TIMSS. The example item for maths class 8 is shown below.



Graph 3.2: The performance of Class 4 students on common questions



Graph 3.3: The performance of Class 8 students on common questions

Class 8

Maths

Subtract:
6000
-2369

A. 4369
B. 3742
C. 3631
D. 3531

Correct Answer: C
Performance of Class 8 Students

Group	% correct
India	59.1
International students (TIMSS)	88.0

Sample Question 58: As per national curriculum framework the concept of 4 digit subtraction with borrowing is introduced at class 5 level. When a question based on this concept was asked in Class 8, Indian students performed lower than international students.

In language, a PIRLS Class 4 passage was asked in SLS 2009 in Class 6 and 8. The passage had 6 questions in total, but international PIRLS performance data was available only for 4 questions. So, a comparison was done for these four questions only. The usual expectation in such cases is that, if a lower class passage/question is asked in a higher class, then student performance would be at least same or higher than the lower class. However, it was found that both class 6 and 8 students have performed lower than the PIRLS class 4 students. Although the performance slightly increased from class 6 to class 8, the class 8 students were much lower than the international average. **This shows that students are not catching up or improving much to the expected level as they move to higher classes.**

Graph 3.4: The performance of class 4 and 6 students on PIRLS questions

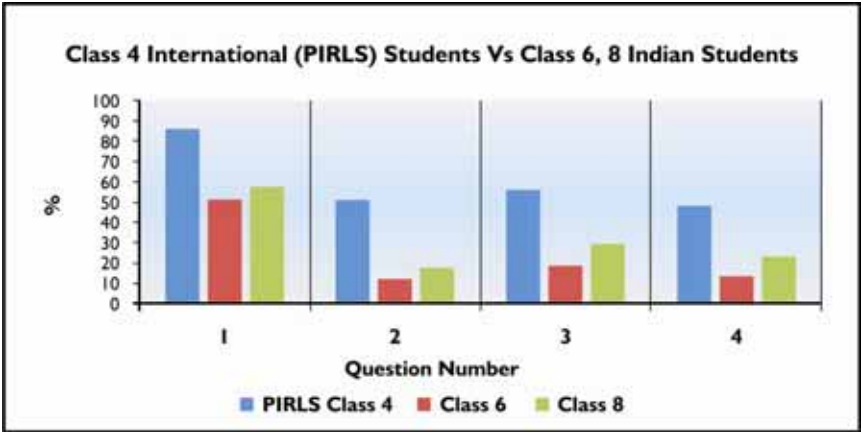
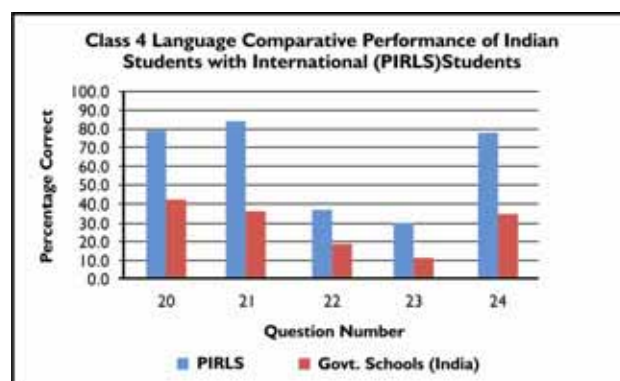


Table 3.6

Graph Q.No	PIRLS Class 4	Class 6		Class 8	
	Students Answering Correctly (%)	Q.No	Students Answering Correctly (%)	Q.No	Students Answering Correctly (%)
1	86.0	30	51.4	43	57.4
2	51.0	31	12.1	44	18.0
3	56.0	33	18.6	46	29.5
4	48.0	34	13.6	47	23.5

In language class 4, five passage questions were commonly asked from PIRLS class 4. In all the 5 questions, Indian students performed much below the international average.

The sample question 60 on the next page is from a passage on how an old man gets rid of mice in his house. The chart on the next page shows the position of Indian students vis-a-vis other nations on the same question.

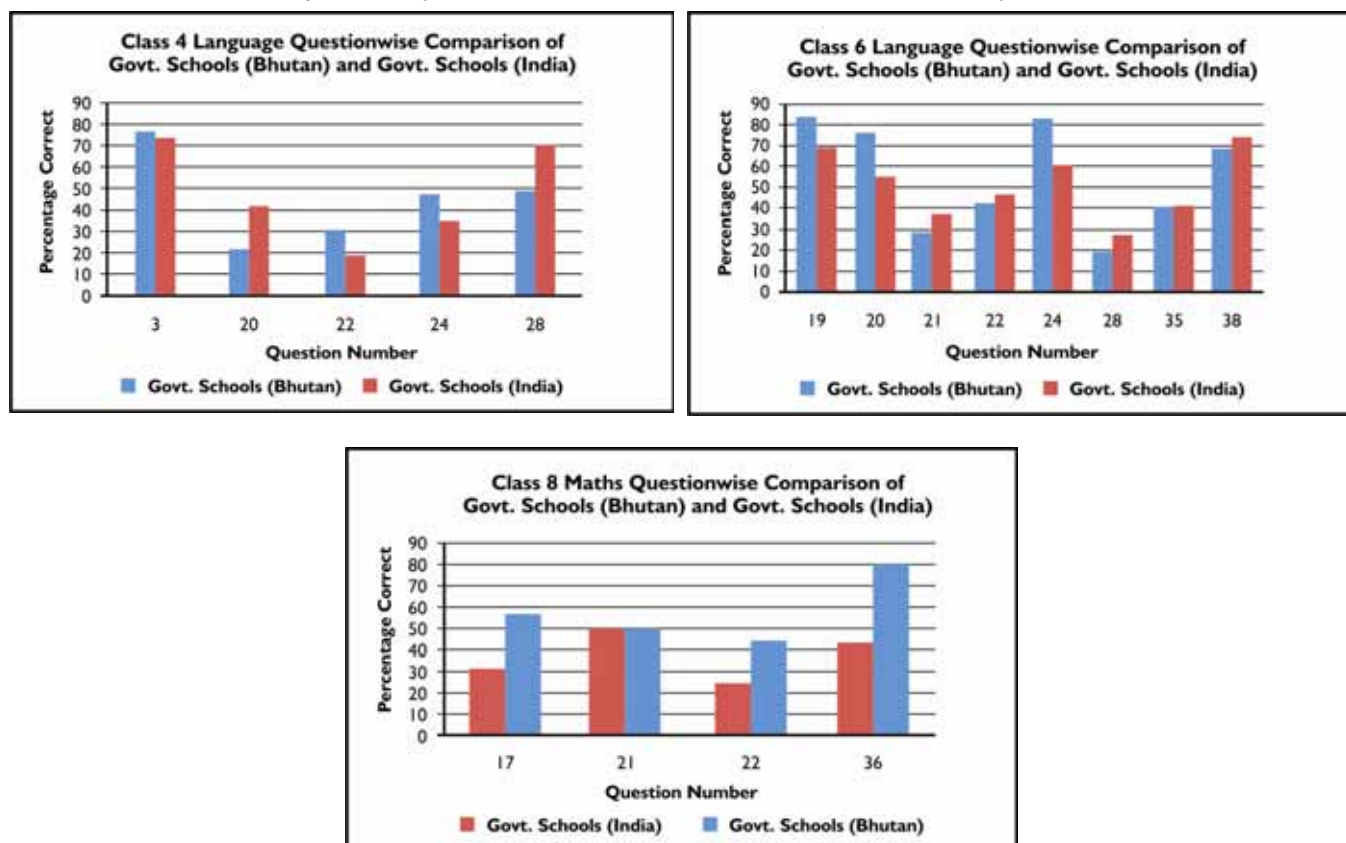


Graph 3.5: The performance of Class 4 students on PIRLS questions

Benchmarking Performance with Government Schools in Bhutan

The performance of Indian students was compared with Bhutanese students in 3 papers – language class 4 and class 6 and maths class 8. This was done to get an idea of student performance in similar school systems in the neighbouring country of Bhutan. In language class 4, Bhutan students are performing better than Indian students in 3 questions but overall performance of Indian students is better. In class 6 Language the Bhutanese students are overall performing better than Indian students. Bhutanese students perform better than Indian students in maths in almost all the questions.

Graph 3.6: The performance of students from India and Bhutan on common questions



Chapter 4. SLS 2009 BENCHMARKS



4.1 What are Benchmarks?

In an approach that moves away from focusing only on the overall achievement scores, this chapter attempts to explain the performances in terms of **what students know and are able to do** at each of the SLS 2009 national benchmarks of achievement – top 10%, upper quarter, median, and lower quarter. As per global practices, the scale anchoring methodology has been followed for this purpose. Items anchoring at each anchor level have been further analysed in order to summarise the knowledge and ability of students on the achievement scale in language and maths.

4.1.1 What is so special about Scale Anchoring?

Classroom experiences, as well as assessment results, clearly show that students perform at different levels of achievement. Students can even be grouped based on their levels of performance. Neither of these, however, provides us concrete steps on what can be *done* to help students performing at lower levels. If we could know firstly, *whether* there are specific topics or concepts which students at lower levels of learning are systematically answering incorrectly, and if so, *what* they are, then remediation could focus on those topics. That is what Scale Anchoring provides, in an extremely scientific manner. The results clearly show which concepts or topics are understood only by students performing at higher levels and which topics are understood by other students also. Certain topics are said to ‘anchor’ at certain percentile levels of performance. Insights provided by the Scale Anchoring process can help us understand the way children learn and plan scientific remediation.

4.1.2 Methodology of Scale Anchoring

Scale anchoring method has been used to summarize and describe student achievement at each of the SLS 2009 national benchmarks¹¹ – top 10%, upper quarter, median, and lower quarter. This meant that several points along a scale were selected as anchor points. The items that students scoring at each anchor point could answer correctly (with a specified probability) were then identified and grouped together. Subject-matter experts reviewed the items that “anchored” at each point and delineated the content knowledge and conceptual understandings each item represented. The item descriptions were then summarized to yield descriptions of what students scoring at the anchor points are likely to know and be able to do.

Scale anchoring is a two-part process. First, the achievement data is analysed for each scale to identify items that students scoring at each anchor point answer correctly. Second, subject matter experts examine the knowledge shown by correct responses to the anchor items, summarise student’s understandings for each anchor point, and select example items to support the descriptions.

An important feature of the scale anchoring method is that it yields descriptions of the knowledge and skills of students reaching certain performance levels on a scale, and that these descriptions reflect demonstrably different accomplishments from point to point. The process entails the delineation of sets of items that students at each anchor point are very likely to answer correctly and that discriminate between performance levels. Criteria are applied to identify the items that are answered correctly by most of the students at the anchor point, but by fewer students at the next lower point.

-‘Describing International Benchmarks’ -TIMSS technical report 1999, page 267

Scale anchoring descriptions in this chapter are based on the national benchmarks, the 90th, 75th, 50th and 25th percentile. The national benchmarks have been computed using the combined data from the schools that participated in the study from all the states. The purpose of scale anchoring was to describe the language and maths that students know and can do at the four national benchmarks.

¹¹ SLS 2009 benchmark refers to the benchmark computed using the combined data from the 18 States and 1UT that participated in the study.

Purpose: Literary Experience

1 Point: Full Credit Sample Response and Results

13. Which words best describe this story?

- ☒ (A) serious and sad
- ☒ (B) scary and exciting
- ☐ (C) funny and clever
- ☐ (D) thrilling and mysterious

Percentage of Students Obtaining Full Credit

Country Average Significantly Higher than International Average		No Statistically Significant Difference Between Country Average and International Average		Country Average Significantly Lower than International Average	
^{2a} Greece	90 (1.5)	Latvia	71 (2.4)	France	63 (2.0)
Cyprus	87 (1.2)	⁺ Scotland	71 (1.9)	^{2b} Israel	61 (1.8)
⁺ Netherlands	87 (1.6)	Hong Kong SAR	70 (1.7)	Macedonia, Rep. of	58 (2.1)
Hungary	83 (1.3)	International AVG.	68 (0.3)	Slovenia	57 (2.1)
Swede	82 (1.1)	Romania	64 (2.3)	Moldova, Rep. of	54 (2.5)
Norway	81 (1.6)			Colombia	52 (2.1)
[†] United States	81 (1.6)			Turkey	47 (2.1)
¹ Lithuania	80 (1.9)			[‡] Morocco	46 (2.4)
Singapore	80 (1.5)			Argentina	45 (2.2)
Czech Republic	80 (1.7)			Belize	38 (1.8)
Germany	79 (1.6)			Iran, Islamic Rep. of	35 (1.6)
New Zealand	77 (2.3)			Kuwait	31 (1.9)
Slovak Republic	77 (1.7)				
^{+2a} England	77 (1.5)				
Iceland	76 (1.5)				
Italy	76 (1.7)				
^{* 1} Canada (O,Q)	74 (1.2)				
Bulgaria	72 (1.8)				
^{2a} Russian Federation	72 (1.7)				
Ontario (Canada)	80 (1.6)			[*] Quebec (Canada)	64 (2.1)

India would
be here

Source: PIRLS, International Report, Chapter 3, Page No. 27

Sample Question 60: For many questions used in the PIRLS test, detailed information is released in the format shown above. This question taken from the PIRLS test was answered correctly by 34.9% of the children in India - this allows us to place the performance of Indian children (in this question) among that of students from other countries.

In order to conduct the data analysis for the scale anchoring, the following steps were used:

1. Anchor points were selected as the 25th, 50th, 75th, 90th percentiles.
2. Group of examinees at each anchor point was formed. For each group, students whose total scores were in between a range of 5 percentile on either side of the target percentile were selected. *For example, for the lower quartile group, students who were between the 23rd and 27th percentiles were selected.* Similarly, the 4 percentile groups were formed.
3. The proportion of students at each anchor points answering the items correctly was calculated.
4. Three criteria were used to identify items for the scale anchoring:
5. A question was said to have “anchored” in a specific percentile group if the students at that level got it right with a high rate (i.e., if 65% of students got it correct) and the students at the immediate lower percentile group level got it correct at a comparatively lower rate (i.e., less than 50% students got it correct). For the 25th percentile, as this is the lowest point, items were checked for 65% of students answering the item correctly.
6. A question was said to have “almost anchored” in a specific percentile group if the students at that level got it right with a high rate (i.e., if 60% of students got it correct) and the students at the immediate lower percentile group level got it correct at a comparatively lower rate (i.e., less than 50% students got it correct). For the 25th percentile, as this is the lowest point, items were checked for 60% of students answering the item correctly.
7. A question was said to have “met 60% criteria” in a specific percentile group if the students at that level got it right with a high rate (i.e., if 60% of students got it correct).
8. The three categories of “anchored”, “almost anchored” and “met 60% criteria” were mutually exclusive and ensured that sufficient pool of items were available at each benchmark.
9. However, in every paper, there are questions which have not anchored in any of the percentile groups. It may be because none of the groups have conclusively got the question correct, i.e. none of the groups have 60% of the students who got that question correct.
10. **The questions which anchor at the 25th percentile are the questions which even the weakest students got correct. Questions which anchor at the 50th percentile are difficult for the weaker students but are grasped by the students slightly better. Questions that anchor at the 75th percentile are got correct by students who have comparatively better ability and questions that anchor at 90th percentile are got correct only by the best students.**

4.1.3 Number of Items Anchoring at Each Anchor Level

Table 4.1

Paper	Low Benchmark 25 th Percentile	Intermediate Benchmark 50 th Percentile	High Benchmark 75 th Percentile	Advanced Benchmark 90 th Percentile	Total Anchored Items	Number of Items in the Paper
Language 4	17	1	2	12	32	37
Language 6	5	9	14	3	31	43
Language 8	7	14	10	7	38	48
Maths 4	14	7	8	8	37	42
Maths 6	2	8	15	12	37	47
Maths 8	4	3	6	8	21	52

4.2 Benchmark Descriptions

What students know and are able to do at the 4 national benchmarks in Maths class 4 is described below. The benchmarks for maths 6, 8 and language 4, 6 and 8 are provided in Appendix C

4.2.1 Maths Class 4 Benchmarks

Advanced Benchmark (students reaching 90th percentile)

90th Percentile IRT Scaled Score: 626

Students understand the concept of place value and know that a 3 digit number can be represented in expanded notation as a sum of the number of hundreds, tens and ones in it. Students understand the concept of multiplication as repeated addition and are able to equate multiplication of 2 numbers as the number of times another number is added. Students know how to form groups of equal numbers out of a given set of an object. Students understand half represented as a fraction and know that it is one out of 2 equal parts and are able to identify the correctly shaded figure based on this. Students can visually identify simple geometrical shapes such as triangles, in a tilted orientation based on their understanding that it is a shape enclosed by 3 straight lines. Students are able to perform conversions between simple units of measurement such as millimetres and metres. Students are able to evaluate and choose the correct mathematical operation to be applied to solve a word problem which requires multiplication of 2 single digit numbers.

High Benchmark (students reaching 75th percentile)

75th Percentile IRT Scaled Score: 568

Students are able to write the number names for 2 digit numbers when their numeral forms are given. Students know the value of single digit whole numbers and are able to identify the greater and lesser numbers. Students are able to identify and count different objects separately and can compare their quantity. Students can perform subtraction of a smaller 2 digit whole number from a larger 2 digit whole number that ends in zero, placed vertically and involving borrowing. Students are able to identify a 3 digit number that is less than another 3 digit number when the difference is given. Students know the arithmetic operation of division and are able to perform simple division of a two digit number that ends in a zero, by a 1 digit number. Students understand parts of a whole and can visually identify equal parts. Students understand the sign of equality and the concept of equations. They are also able to apply simple arithmetic operations in solving equations.

Intermediate Benchmark (students reaching 50th percentile)

50th Percentile IRT Scaled Score: 504

Students are able to write the numeral form of 3 digit numbers when their number names are given. Students can recognise missing numbers in an ascending sequence of 2 digit numbers greater than 20. Students can perform multiplication of a 2 digit number with a 1 digit number, not involving carry over. Students understand the concept of half as a number divided by 2, for example they know that in a group of same objects, 3 out of 6 is half the number of that object. Students are able to recognize patterns in sequences involving shapes. Students are able to apply subtraction operations appropriately and successively in word problems. They are able to represent 'nothing' that remains after the subtractions as a 'zero'. Students are able to apply the appropriate mathematical operation in word problems involving day-to-day life situations such as finding the age of a person from that of another person when the difference is known.

Low Benchmark (students reaching 25th percentile)

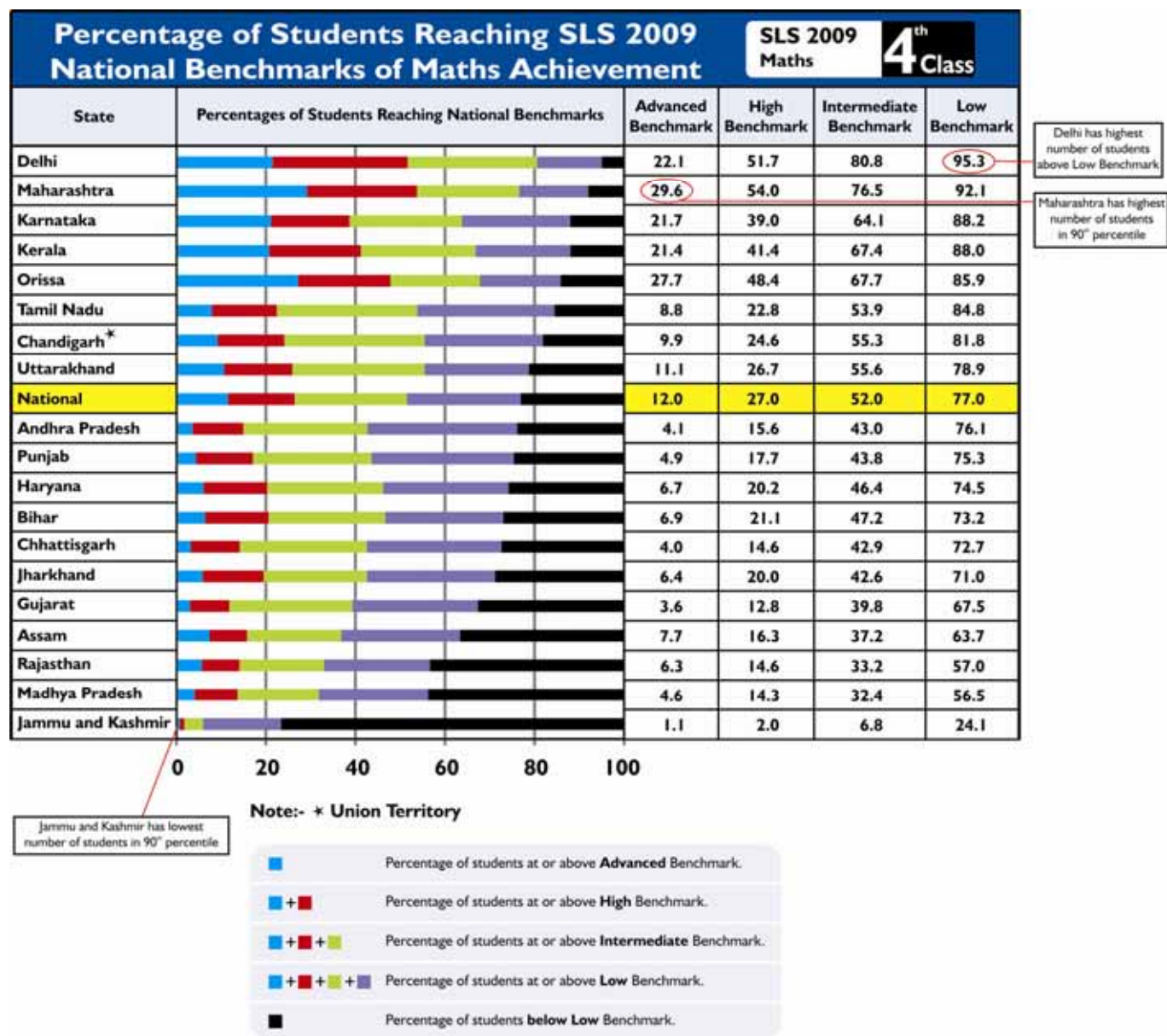
25th Percentile IRT Scaled Score: 439

Students write the numeral form of 2 digit numbers when their number names are given. Students know 2 digit numbers less than 20 and can identify a number that is missing from a sequence of consecutive numbers. Students complete the missing numbers in a descending sequence of 2 digit numbers less than 20. Students know the arithmetic operations of addition and subtraction and are able to carry them out for 1 digit numbers, placed vertically. Students do addition of two 2 digit numbers, placed vertically and involving carry over. Students do subtraction of 1 digit numbers from 2 digit numbers, placed vertically and not involving borrowing. Students know the arithmetic operation of multiplication and are able to perform the same for 1 digit numbers, placed horizontally. Students identify similarities and differences visually, between different 2D shapes. Students understand fractional quantities such as half written in a word form as one out of 2 parts and apply them practically in their daily context. Students read time from a traditional analog clock. Students are familiar with the traditional currency denominations and can perform simple addition of the same. Students can read the price tags of objects whose values are represented by 3 digit numbers, to find the one which costs the least. Students understand word problems based on counting the number of objects present in loose and in bundles of ten.

4.3 How Different States achieve the Benchmarks

The graph below describes how different states are reaching the different benchmarks in Class 4 Maths. Similar benchmark graphs for maths 6, 8 and language 4, 6 and 8 are provided in Appendix C

The states are sorted/arranged in the order based on the number of students scoring at or above 25th percentile (low benchmark) nationally. The section of each bar that is coloured black indicates the number of students who fail to reach even the low benchmark.



Jammu and Kashmir, Madhya Pradesh and Rajasthan have the maximum number of students who fail to reach the low benchmark, as could be seen from the extent of the black coloured section of their bars.

Maharashtra has the maximum number of students (29.6%) reaching the Advanced Benchmark (90th percentile nationally); Jammu and Kashmir, Gujarat, Chhattisgarh and Andhra Pradesh have a negligible number of students (1.1%, 3.6%, 4.0% and 4.1% respectively) in the Advanced Benchmark.

At the national level 33.0% of students do not reach the low benchmark and about 12.0% of students reach the advanced benchmark.

4.3.1 How is this information useful?

Benchmark information can be used in a variety of ways

- At the policy level
- At the level of curriculum and pedagogy to determine patterns and in-explicables
- At the school level

I. At the Policy Level: Educational reform is a complex process and requires focussed and consistent approach to bring about a change. Often the various stakeholders are at a loss where to prioritise their time and resources for lack of availability of data. For example, **a policy maker deciding about the training required for teacher education, may need information on which subjects and concepts need the most attention.** The policy maker would like to know if training needs to be provided for all concepts or is there a way in which the concepts requiring the most attention for a state/district can be identified and targeted. Let us see a few cases of how the information can be used.

CASE I: Benchmarking determines what students know and are able to do at different quartiles of the population. Benchmarking also makes it possible to get this type of specific and accurate information about the different segments of student population. For example, in the graph above, **Let us take the case of Madhya Pradesh.** In Madhya Pradesh while 56.5% reach the low benchmark the remaining 43.5% are not acquiring the knowledge and skills specified by the low benchmark. For instance, these 43.5% students in Madhya Pradesh in class 4 are NOT able to do the question below which a student reaching the low benchmark can do.

An Example of What Students Reaching Low Benchmark In Maths Class 4 Knows and Are Able To Do

Class 4	Maths
Write the number that comes between: 18, _____, 20	Students' reaching the low benchmark knows 2 digit numbers less than 20 and are able to identify a number that is missing from a sequence of consecutive numbers.
<i>Sample Question 59 Majority (86.9%) students are able to give the correct answer. A big number of students are able to answer this question as it is anchored at low benchmark (students reaching 25th percentile). (Ref: Annexure C)</i>	

Madhya Pradesh has also only 4.6% of students reaching the advance benchmark. For instance, only 4.6% of students in Madhya Pradesh can DO the below question.

An Example of What a Student Reaching Advanced Benchmark In Maths Class 4 Knows and Are Able To Do

Class 4	Maths
Distribute the bananas below into groups of four. How many groups will be formed? Tick (✓) the answer.  A. 12 B. 6 C. 4 D. 3	A student reaching the advanced benchmark knows how to form groups of equal numbers out of a given set of an object. This is not known to the students reaching the high benchmark.
<i>Sample Question 60 A big percentage of students (37.9%) are choosing option A. It seems these students are simply counting the number of bananas instead of grouping them. The ability to group objects in equal numbers from a given set is acquired by higher ability students.</i>	

While this information is alarming, it DOES assist a policy maker to take data driven decisions. The policy maker in this case is confronted with the issues that at the lower end of ability, there are many (43.5%) who are not reaching the most basic education, while at the higher end, there is a paucity (4.6%), with students not reaching their full potential. The policy maker

can decide to prioritise on the lower end or the higher end. If the State decides to focus its attention on improving the 43.5% of students below the low benchmarks – then most of the efforts of the state, say remedial education, teacher training, textbook reform/assistance, etc for this class and subject should focus on the knowledge and skills that the low benchmark students achieve, rather than trying to improve ALL of the system in one go. This way one can assist the students below low benchmark first to acquire the knowledge and skills that are needed to reach the low benchmark; then students achieving low benchmark to acquire the knowledge and skills defined by the intermediate benchmark and so on.

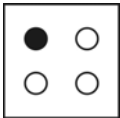
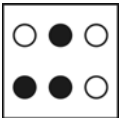
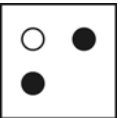
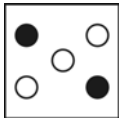
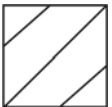
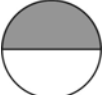
CASE 2: Similarly, using the graph above, Tamil Nadu may find that in comparison with Orissa it has almost similar number of students who have crossed the low benchmark (Orissa – 85.9%, Tamil Nadu – 84.8%), however, it has a low number of students in the Advanced Benchmark in comparison to Orissa (Orissa – 27.7%, Tamil Nadu – 8.8%). Tamil Nadu may then decide to target improving the number of students reaching the advanced benchmark. The state education officers as a result may focus their efforts on bringing the ‘High’ benchmark students (students not reaching the advanced benchmark) to the advanced benchmark level by providing additional learning support for the concepts that form the advanced benchmarks.

Thus at the policy level, knowing the benchmarks and what percentage of students reach these benchmarks allows a decision maker to set their goals and focus on specific target areas rather than widely spread their attention, thereby spreading themselves thin.

II. At the level of curriculum and pedagogy to determine patterns: The curriculum and pedagogy efforts can easily provide remedy for what is not known or understood by the different ability groups by observing the benchmark patterns across classes.

CASE 1: The same concept asked in different levels of depth through different questions, often, through scale anchoring provides information on how students at different ability levels are able to achieve them. For example, all the questions below that anchored at the different ability levels of Class 4 – Advanced, High, Intermediate and Low explore the concept of a fraction as understood by students.

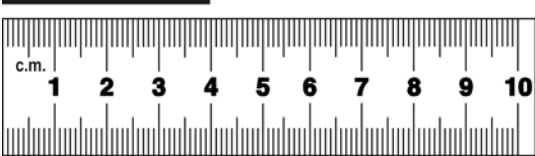
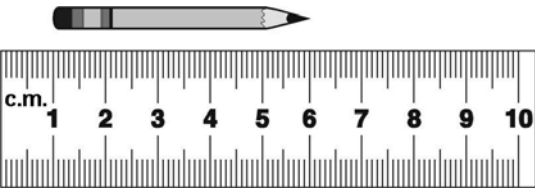
Class-4 Maths

Concept of a Fraction		
Low Benchmark (25 th Percentile)	If a watermelon weighs 10 kg, how much will half the watermelon weigh? _____	Students understand fractional quantities such as half written in a word form as one out of 2 parts and apply them practically in their daily context.
Intermediate Benchmark (50 th Percentile)	In which figure are one-half of the dots black? Tick (✓) the answer.  A.  B.  C.  D.	Students understand the concept of half as a number divided by 2 , for example, in a group of same objects, they know that 3 out of 6 is half the number of that object.
High Benchmark (75 th Percentile)	Which figure is divided into four EQUAL parts? Tick (✓) the answer.  A.  B.  C.  D.	Students understand parts of a whole and can visually identify equal parts .
Advanced Benchmark (90 th Percentile)	Which figure is $\frac{1}{2}$ shaded? Tick (✓) the answer.  A.  B.  C.  D.	Students understand half represented as a fraction and understands it as one out of 2 equal parts and are able to identify the correctly shaded figure based on this.

It is interesting to note that while the low scoring students (25th and 50th percentile) are able to understand a fractional half as one out of two OR its equivalents (3 out of 6), they have difficulty in understanding equality of parts and does not know that all the parts have to be equal.

CASE 2: *Benchmarking reveals that some concepts although expected to be achieved in class 4, are achieved only by the brightest students in class 8.* For example, the question below on measurement was asked in the class 4, 6 as well as class 8 Maths paper, but was achieved only by students reaching the advanced benchmark in class 8.

Class 8 Maths

Measurement of Length		
Advanced Benchmark (90th Percentile)		Students understand information presented visually and are able to calculate distance by observing the starting and the ending points and calculating the units of distance travelled between the 2 points. Students can apply their knowledge of measurement to find out the length of a pencil's image using the image of a scale.
	The length of the line in the figure above is 4 cm. How long is the pencil shown in the picture below? (Use the ruler shown in the picture.)  _____ cm.	

CASE 3: *Benchmarking reveals that certain concepts which are difficult and not achievable except by the brightest students in a lower class continue to remain difficult in the higher classes too.* It is natural to expect that while a concept is difficult for a class 4 student, it will become comparatively easy to a class 6 student and much easier to a class 8 student. In such cases, one would see a question repeated across classes to anchor say at 90th percentile in class 4, then at 75th percentile at class 6 and at 50th percentile in class 8. However, benchmarking reveals that certain concepts which are difficult and not achievable except by the brightest students in a lower class continue to remain difficult in the higher classes too, indicating that these will largely remain in the future adult population too unless remedied. For example, the concept of multiplication below remain achievable only by the students at the 90th percentile in both classes 4 and 6, revealing that the concept has not become easier at class 6 as one would expect.

Class 4, 6 Maths

Concept of Multiplication		
Advanced Benchmark (90th Percentile)	Fill in the appropriate number in the box. 3 × = 3 + 3 + 3 + 3	Students understand the concept of multiplication as repeated addition and are able to equate multiplication of 2 numbers as the number of times another number is added.

Similarly in language, writing a few meaningful sentences correctly seems to be the most difficult and is achieved only by the students who reach the high or advanced benchmark. The questions below, asked in class 6 and 8 to write 3 sentences about the given context, show that this remained difficult. They were only answered by achieved by only students in High benchmark in both class 6 and class 8 and had not become easier in class 8.

Writing		
High Benchmark (75 th Percentile) (Class 6 Question)	Write 3 sentences about yourself. _____ _____ _____	Students are able to write 3 meaningful sentences appropriate to the context. Students are able to construct these sentences applying rules of grammar correctly.
High Benchmark (75 th Percentile) (Class 8 Question)	Write 3 sentences about your favourite game. _____ _____ _____	

III. At the school level: A teacher can administer or include the sample questions that are released for the different benchmarks in her class tests. This will enable the teacher to check if the students in her class are reaching that benchmark or not. The sample items and the questions that anchor at different levels in language and maths for classes 4, 6 and 8 are given in Appendix C. An example is shown below.

SLS 2009 High National Benchmark (582) of Language Achievement - Example Item		
An Item That Students Reaching the High National Benchmark Are Likely to Answer Correctly*		SLS 2009 Language 4 th Class
Competency: Reads short text of 5-6 sentences that describes daily activity, routine context, simple description, simple story independently and comprehends stated facts Description: Understanding stories that are read out and answering questions that require drawing straight forward inferences from explicit information To answer this item correctly, students need to connect the two events the laying of eggs and the baby crocodiles coming out of the eggs. While three months are explicitly mentioned, in connection with hatching, it is only implicitly connected to the laying of eggs.	State	% Correct
	Delhi	75.2 ▲
	Kerala	71.8 ▲
	Maharashtra	70.2 ▲
	Orissa	69.2 ▲
	Karnataka	66.2 ▲
	Tamil Nadu	62.5 ▲
	Chandigarh *	58.2 ▲
	Punjab	57.5 ▲
	Uttarakhand	57.2 ▲
	National	56.6
	Bihar	55.0 ▼
	Andhra Pradesh	51.9 ▼
	Assam	51.3 ▼
	Chhattisgarh	49.8 ▼
	Gujarat	48.2 ▼
	Jharkhand	47.5 ▼
	Haryana	46.9 ▼
	Rajasthan	41.1 ▼
	Madhya Pradesh	38.5 ▼
	Jammu and Kashmir	21.8 ▼

▲ State average higher than national average.

▼ State average lower than national average.

EXCERPT

Crocodiles lay their eggs and cover them with sand. After three months the baby crocodiles are ready to come out of the eggs.

When do the baby crocodiles come out of the eggs?

- A. one months after the eggs are laid
- B. two months after the eggs are laid
- C. three months after the eggs are laid ✓
- D. four months after the eggs are laid

Chapter 5. RECOMMENDATIONS



5.1 Overall Recommendations and Policy Suggestions

It is worthwhile to review the main findings briefly here and then discuss what they suggest about the current educational scenario in the public schools in the country. The main findings of the study are:

1. Learning levels are extremely low. In the lower classes, a fair amount of 'rote-based' or 'procedural' learning is evident in very basic numeracy skills such as number sequencing, operations of whole numbers involving 1 or 2 digits; naming of numbers; reading clock time, understanding currency, etc. Even among procedural questions; students are able to comparatively handle only 'straightforward' questions that are closer to what one would practise from a typical textbook and not when they are slightly atypical. In higher classes, students are falling behind in all learning, even procedural. Responses to some questions suggest that students are probably coping through learning happening outside the class. For example, in class 4, while more students could add the fractions $2\frac{1}{2}$ and $1\frac{1}{2}$ in a word problem using a real life context, fewer of them could add the same when asked as a straightforward addition question as ' $2\frac{1}{2} + 1\frac{1}{2} = \underline{\quad}$ '.
2. Learning taking place is not 'Learning with understanding' and a number of misconceptions exist among students on the concepts learnt. The learning that is happening seems to be procedural or rote-based and not one of 'Learning with Understanding' as students find it difficult to answer questions that require a deeper understanding of the concept.
3. Learning gains seen across classes is slightly incremental and not a large jump. Student performance in common questions that were used to check learning gains across classes showed that performance increased as students move from class 4 to 6 to 8 in both language and maths. However, the extent of improvement was often slightly incremental and not a large jump as one would expect. In most cases, nearly 40% of students in class 8 do not seem to have acquired class 4 competencies. Many misconceptions that students have in lower classes still continue in higher classes and in some cases were found to even become stronger.
4. Students find it difficult to express their thoughts in their own words in writing. Their writing does not go beyond the most basic, tried and tested formulaic sentences they probably trained for while in their lower classes. Their writing shows that they are learning language more as a subject and less as a means of natural communication. The writing has a number of errors in spelling and grammar, and punctuation marks are conspicuous by their absence.
5. In all the states tested, fewer students were found to comprehend what they read. For e.g., in the oral reading test, more than 87.1% of students in Gujarat could read a simple word, but only 40.3% could read a short passage and only 22.6% could comprehend the information implicit in the passage they read.
6. There are significant state-wise differences in student performance. Based on the relative performance/rank of the states in different classes and subjects, an attempt has been made to consolidate the performance of 17 states. Kerala, Maharashtra, Orissa and Karnataka are clearly performing overall better than the national average. Jammu and Kashmir, Madhya Pradesh and Rajasthan were among the states that ranked among the bottom three overall. Bihar performed the same as national average.
7. Extent of students scoring zero and the overall performance of a state showed differences. Andhra Pradesh, Assam, Chhattisgarh and Jharkhand, although below the national average, have fewer students not scoring in the test compared to better performing states such as Haryana and Karnataka, indicating that while the states' efforts could be addressing the lowest ability students, it does not provide adequate support for overall improvement of all students. Jammu and Kashmir, Madhya Pradesh and Rajasthan which ranked among the bottom 3 in overall performance also had the highest number of students scoring zero indicating that these students are being left behind in these states.
8. The levels of learning of Indian students in government schools in class 4 and class 8 tested is much lower than the international average as represented by studies like Trends in International Maths and Science Study (TIMSS) and Progress in International Reading and Literacy Study (PIRLS). Similarly, on common questions used from an extensive study for private schools¹² in India catering to the elite and upper middle class, students in government schools showed a much lower performance.
9. The comparative performance of boys and girls is similar to many international studies – boys seem to do better in maths, although the SLS study reveals that it is a meaningful difference that matters in class 8 only.

¹² Questions were taken from ASSET, a diagnostic assessment test by EI, in which more than 4 lakh students participate from all states of India

10. Analysis of Background factors showed that when students perceive themselves as being good at studies and think of school as a place of fun and learning, they show good performance. The analysis also reveals the importance of inculcating reading habits in student, for spending 30 minutes each day to read material other than textbooks brings about higher achievement. Teacher training is important especially for mathematics teachers for better achievement of students in the subject. Unexpectedly, analysis shows that students whose teachers have an academic degree developed lower scores in both language and mathematics than those who did not have these.

5.1.1 Discussion on the Findings

The study emphasizes that the levels of learning in our Indian public schools is low. The problem appears to be two fold – very basic or rote learning in the lower classes and not much progress even in rote learning in higher classes. Students in higher classes are also not improving in their lower class competencies, as is seen from questions that were used commonly across classes to check for progression in learning. The state policies of promoting students, irrespective of their academic performance may be moving them up in class level but are not doing so in actual learning.

In the last decade, while there has been considerable progress in enrollment to achieve near 100% enrollment of 6-14 year olds, this is only the first step. The children must also complete eight years of useful and relevant school education (11th plan document, GOI, 2006). In the 21st century we need students who are learning the portable skills – of critical thinking, creativity, learning to learn and adapt to new ways as the world changes. While providing access and inclusive education to all is important, it is critical that steps are taken to ensure that this education is one that is equipping the student with the needed skills and competencies.

Unless educational systems have been inherently focussed on students' learning with understanding, it should not be surprising that a critical comparison usually discovers that student learning levels are low. This has been the experience the world over - and historically, educational systems were not focussed on student learning. Reforms and changes have mostly been initiated in the last 20 years, with some nations (like Finland and Singapore) showing significant success. India can, and should, aspire to be performing at the levels of those countries within a period of 5-10 years. The key requirement to achieve this is political will and a shared understanding of the goal and what is needed to achieve this goal.

If the goal is to improve the quality of an educational system, we need to define quality clearly and how we would know if it is achieved. Early definitions of quality focussed on the physical infrastructure and teacher qualifications. Later the quality of classroom processes was considered a measure. However, from the viewpoint of *outcome-focus* and *objectivity*, it is widely agreed these days that measuring student learning levels is the most effective way to judging the overall quality of the school education system. This does not mean that factors like physical resources and classroom processes are not important, rather those are factors that determine quality. However, to know that these factors are serving the goal well, the focus needs to be on the outcome '*how well are students learning?*' which needs to be regularly monitored. Scientifically collected data on student learning outcomes and the existing gaps, then, enables the stakeholders in the system to address the gaps to attain the quality goal.

5.1.2 Strengths and Weaknesses of the Existing System

Looking at the system holistically, we feel the strengths and weaknesses of the public education system are as follows:

Strengths:

1. **Fairly Qualified Teachers:** Overall teachers in government schools are more likely to be trained, have greater experience and a higher salary than teachers in private schools. Regular teachers are also more educated than teachers in the private sector (Sangeeta Goyal and Priyanka Pandey, 2009). We also found that teachers in the government schools tested were fairly well qualified. (*For example, out of the 2399 schools tested in the study – 75.0 % of teachers have an academic degree or more and only 4.2% of teachers reported not having any teacher training*)
2. **Education is valued in the society:** All sectors of the society especially the middle class and the lower class value education as a gateway for a better life. The success of the ordinary citizens to become wealthy in sectors like IT also serves as an inspiration for the lower classes to send their children to school. The right to education act also shows a society that values education as a fundamental right. (*All this is reflected in the opinion of students too, only 16.0% of students tested said that they considered coming to school as useless*).

What are Learning Outcomes?

Learning outcomes are statements of what a learner is expected to know, understand and be able to demonstrate after completion of learning. They are essentially student centred or learner centred. They seek to describe the student's learning progress in terms of the knowledge acquired, the comprehension of that knowledge, the capacity to apply it, the capacity to analyse, synthesise and evaluate. Learning outcomes guide the selection and coordination of appropriate content, learning activities, and assessment strategies that promote the overall learning process. Quality of student learning can be monitored against the expected performance for these learning outcomes.

An example of a Learning Outcome in Reading: Students' listen to or read the various types of texts for information, comprehension, and literary appreciation.

In Classes K-4, students should know and be able to do the following:

- Listening to /and reading a variety of Indian and non-Indian literary (class-appropriate fiction, folktales, fables, funny stories, rhymes, plays, diaries, etc.) and non-literary text (simple informational text, text books of other subjects, picture books, simple descriptive/ narrative text, children's encyclopaedia, instructions from science experiments, factual recounts (news stories), lists, etc.

- From 'Learning Standards' by Educational Initiatives

How can one check for the quality of attainment in this reading outcome?

These are described in the 4 performance levels – Basic, Partially Proficient, Proficient, Advanced.

Basic Level

This requires students to receive or recite facts or to use simple skills or abilities. Oral reading that does not include analysis of the text, as well as basic comprehension of a text, is included. Items require only a shallow understanding of the text presented and often consist of verbatim recall from text, slight paraphrasing of specific details from the text, or simple understanding of a single word or phrase. Some examples that represent, but do not constitute all of, Basic Level performance are:

- Support ideas by reference to verbatim or only slightly paraphrased details from the text.
- Use a dictionary to find the meanings of words.

Partially Proficient Level

This level includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Inter-sentence analysis of inference is required. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed. Items require closer understanding of text, possibly through the item's paraphrasing of both the question and the answer. Some examples that represent, but do not constitute all of, Partially Proficient Level performance are:

- Use context cues to identify the meaning of unfamiliar words, phrases, and expressions that could otherwise have multiple meanings.
- Predict a logical outcome based on information in a reading selection.
- Identify and summarize the major events in a narrative.

Proficient Level

At this level, students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts. Some examples that represent, but do not constitute all of, Proficient Level performance are:

- Explain or recognize how the author's purpose affects the interpretation of a reading selection.
- Analyze and describe the characteristics of various types of literature

Advanced Level

Higher-order thinking is central and knowledge is deep at this level. Students take information from at least one passage of a text and are asked to apply this information to a new task. They may also be asked to develop hypotheses and perform complex analyses of the connections among texts. Some examples that represent, but do not constitute all of, performance at advanced level are:

- Analyze and synthesize information from multiple sources.
- Examine and explain alternative perspectives across a variety of sources.
- Describe and illustrate how common themes are found across texts from different cultures.

- Dept of Education, Colorado

3. **Learning in mother tongue:** A number of research articles point out that learning in mother tongue is beneficial in the initial years of schooling. The government school system is the largest and often probably the only school system in the different states of India that offers education with the mother tongue being a medium of instruction.
4. **Ability of the State to handle large programs:** The state has the machinery well entrenched to handle large programs, the scale of which cannot be compared to any other sector. (*For example, the midday meals program is the largest school lunch programme in the world and caters to about 120 million students*). This is important as any state program that is adequately planned and meaningfully designed can be implemented and bring about a larger change more easily compared to any other sector.

Weaknesses:

1. **Focus on rote learning:** The study reveals that whatever learning taking place is extremely rote-based, focussing on the most basic facts, recall and procedures and less on thinking, understanding and helping students express their own thoughts.
2. **Students lagging much behind in higher classes:** The study also shows that as students are moving up the classes, there is a fall in their performance in all learning including rote. This means that students may move up the classes based on social promotion. Their physical presence in schools in these classes is not guaranteeing learning.
3. **Focus on 'inputs' rather than 'outcomes':** While there has been a number of steps taken to provide inputs such as physical infrastructure, classroom processes, etc the system so far has not much focussed on outcomes – 'are children learning?' – focus on outcomes maximises efficiencies in any system. When outcomes are kept as the goal, the monitoring and providing of inputs is well integrated to achieve the goal.
4. **No empirical research to drive system shift:** There is paucity of empirical research to establish the needs of the system in terms of student learning, teaching learning process, etc. The decision makers often have to base the decisions on ad-hoc compulsions as data that provides evidence and is collected scientifically to provide insights is often not available.
5. **Too much diversity/ tiers of schools:** Government schools are run usually for the very poor, the poor go to low-cost private schools, the urban middle class to private schools, while the very rich go to exclusive elite schools. There is no positive public pressure on the schools, especially the ones that cater to the very poor to perform.
6. **Accountability/Motivation:** Teacher absenteeism study done by Harvard (Kremer et al, 2004) showed that nearly 25% of teachers were absent from school, and only about half were teaching, during unannounced visits to a nationally representative sample of government primary schools in India. The current system does not have a clear performance metric to differentiate a teacher who is performing and one who does not.

Based on this analysis of strengths and weaknesses, we make the following recommendations. Some of these can be affected through policy changes. Others however, need building public opinion which is a gradual process and some are about setting in motion certain regular practices that will help in this movement.

5.1.3 Recommendations

1. **Orienting policies to focus on learning outcomes:** While the Central Government initiated Sarva Shiksha Abhiyan (SSA) in 2002 and set the targets of universal primary education in 2007 and Universal Elementary Education (UEE) by 2010 respectively, it has not set itself a target based on quality of student learning outcomes. Hence the expected 85% transition rate from Class VIII to IX¹³, as revealed by the study may well be students who are ill equipped with the knowledge and skills expected for that class. While a debate may rage on "what is quality education?" and on the importance of imparting moral values, student attitudes and other co-curricular activities, one cannot deny that basic education needs to ensure that the key concepts and skills are attained through formal schooling. The deficiencies in the other non cognitive areas can be adequately compensated even by society, but school is the main centre where the student will have access to develop the competencies and skills expected of his educational level. It is important that all policies are build with a focus on the learning outcomes, for, if students are not learning, then this defeats the purpose even if all other goals are met.
2. **Making Low Stakes Diagnostic Assessments – a regular feature of the state programs:** The assessments should not be done with a view to penalising or punishing low-performing students, teachers or schools. Many systems, including the well-known US No Child Left Behind policy, penalise low performing schools or teachers or de-recognise them. Such measures create fear in the system and lead to manipulations including outright cheating. They can seriously undermine

¹³ Report of the CAGE committee on "Universalisation of Secondary Education", GOI, 2005

the original objective of trying to *improve the system*. Assessments tend to work well when they are low-stake. **Their purpose is to inform students, teachers, schools and even the larger society where schools stand. Without official pressure**, the purpose of the assessment is two-fold: to provide support and information, and this itself leads to the creation of a positive peer and/or self pressure. Teachers are forced to ask themselves why their students in class 6 (say) are not learning well when other children in another part of the district or state are. Punitive action – if at all – should remain a last and rare option.

It is sometimes felt that tests developed and administered by the class teachers themselves can be used. But such tests and processes contain conscious and unconscious biases. The tests must be created by a body or organisation that is truly independent and is not in way responsible for providing quality education. It should have a standing in the research community and must have an established research track record (including benchmark data) to develop credible assessments. Low stakes assessments thus developed should be diagnostic and check for understanding and not simply rote or recall.

3. **Extensively use Benchmark data from the study to build reform:** A powerful goal of these assessments is to obtain a clear picture on where students and teachers stand with respect to peers in the state, peers in the rest of the country and peers internationally. These are not just numbers but detailed statements of strengths and weaknesses which can lead to specific action points. This is one important and powerful positive consequence of tests such as SLS. A second is simply the motivational value – to know that others are doing something better provides a challenge and a certain amount of pressure which can be a powerful motivator. (Point 2 above emphasises that additional pressure should not be applied). Note that benchmarking also reveals one's strengths, not just weaknesses and this can be a positive motivator.

Benchmarking can and often does reveal overall weaknesses and that the state of affairs is not as rosy as it was believed. Elements like the media and others would try and exploit this, and that is one reason governments are often sensitive about such results (this is true across the world). However, the courage to openly accept that there are weaknesses that need to be improved and to face the criticism may be an important step towards focussing on the problems and solving them.

This is an excerpt¹⁴ from a press note released after Dubai's results in an international benchmarking test (which were not very good) were revealed: *"The head of Dubai's Knowledge and Human Development Authority said: 'You cannot improve your education system overnight. Neither can we wait 12 years to improve our system, since that is the school life of a child and their education is too important to usWe want to share information with the public so that everyone is working with us as we go forward with our job of developing the human potential needed for Dubai's continuing success. We need all of our human resources to collaborate on improving education.... Although Dubai's results have placed us first in the Arab world, we do not intend to compare ourselves with countries in the region. Our commitment is to found an education system of international standing.' The CEO of Dubai Schools Agency, added: 'Now we can move on, taking our teachers with us. We have all had to confront the disruption and anxiety that assessments can bring. We know that good education is all about good teachers, and we can now analyse and develop data which will help us in supporting our teachers. For example, perhaps teachers spend too much time on certain subjects and not enough on others. We have real evidence now and we can use it to make adjustments if they are necessary.'*

4. **Student Progress Tracking System:** Using data to make informed decision-making at every level changes how institutions have to function and it requires changing expectations, changing the thinking of management at the leadership level, developing assessments that accurately measure learning gains, developing information technology systems that can make this data available accurately and in a timely way, and building capacity at every level of schooling to use data well.

We have reached a stage in our development where accuracy of available data alone can make a significant difference to our development. This brings to the fore the need for computer-based systems like Student Progress Tracking System in helping to achieving universal primary education nationally. Student Progress Tracking System is a computerised system with unique student identification and aids in tracking student learning / progress on both scholastic and non-scholastic domains. The system would be a comprehensive database of students, teachers, schools and is ideally accessible at the school as well as at a central level. Such a system will allow drilling down of information to the level of the individual student and teacher and enable targeting the remediation.

¹⁴ <http://www.khda.gov.ae/CMS/WebParts/TextEditor/Documents/TIMSS%20English%20Press%20Release.pdf>

- 5. Large Scale Awareness campaign to redefine attitudes towards learning - a movement *against rote learning and for Learning with Understanding*:** Rote learning can deceptively look like learning and be mistaken for it. A consensus needs to be gradually built that *rote learning is not learning at all*. Further, our entire education system is geared to measure rote and this is well entrenched in the exams at the school and board level. It can be argued that ‘reports of low student learning levels are alarmist, as students are scoring well in board exams and securing college admissions, and so, outside comparisons are unfair because the ground realities in our country are unique’. Such an argument misses the point that the world of tomorrow needs people who may need to learn new skills every few years, and rote learning will not serve them well. Much debate and differences about education quality stem from whether the goal being sought is rote learning, or learning with understanding. A consensus needs to be gradually built that *rote learning is not learning at all*.

A long term public education campaign should also be instituted with the idea of discussing and disseminating these views. Discussions and campaigns should be arranged around issues like “What is learning?”, “What is good education for our children?”, “What is a good school and how can I help improve my child’s school?”, etc.

This recommendation sounds general and possibly a little vague, but is an important one. Teachers and parents need to know what success will look like – a rote learning system says success means high marks in the board exams – a system based on learning with understanding says success means students pursuing careers of interest with passion, a culture of entrepreneurship and an education system that genuinely engages learners. These two goals are very different, and through discussions and debate, one of them has to be agreed upon, and then pursued with full vigour.

- 6. Widely disseminating the findings of this report among teachers and others:** We recommend that each state make a systemic and detailed plan to disseminate this report, giving every teacher an opportunity to see it, understand and even question and discuss it. We recommend that a series of workshops be held at the district level in each state to share the findings with teachers, and other teachers be provided simplified copies of this report. The purpose of all this is two-fold: 1. initiate the rote versus learning with understanding debate in the country; and 2. get teachers to start thinking about ‘what are children learning and not learning’ rather than just focussing on ‘what are we teaching or what to teach’.
- 7. Providing effective teacher support based on the feedback from the assessment:** Targeted capacity building in teachers is often enabled by detailed diagnostic assessments such as the SLS, which reveal absolute performance and trends showing weaknesses in groups of students, schools and sometimes even a region or in the entire state. The background factor analysis in the study reveals that teachers with academic degree were ineffective in achieving high scores for their students in language and maths while teachers with teacher training qualification were effective in achieving higher scores for their students in maths. This is an indication for reviewing the teacher recruitment and training system.
- The key goal of teacher training has to address teacher attitudes towards learning, specifically on rote learning versus learning with understanding
 - Student interviews¹⁵ should become a standard part of pre-service teacher training, and also an activity that practising teachers should be encouraged to do. Practices like Japan’s ‘Lesson Study’ which make a teacher more reflective and research-oriented, even in a classroom context, should be analysed and adopted.

5.2 Limitations

The issue of student seriousness: The reliability of data collected in tests like these is sometimes objected to on the grounds that students may not be performing to the best of their abilities in low-stake tests like these. According to this argument, it is only when students recognise that the cost of not performing ‘well’ is high, that they do their best. Hence, if this is true, the results may be lower than what students are really capable of. However, the relative performance will still be comparable with international and Indian students as the same argument holds true for all low-stakes tests.

Testing limited to language and maths: One may argue that concluding about educational quality on the basis of performance in language and maths alone is one dimensional, while quality education is much more and includes the non-cognitive – moral and social aspects too. However for the purpose of measurement, it has been generally accepted through research that acquisition of cognitive skills is a reliable indicator of attainment in non-cognitive areas too (Govinda and Varghese, 1993, Hanushek, 2007).

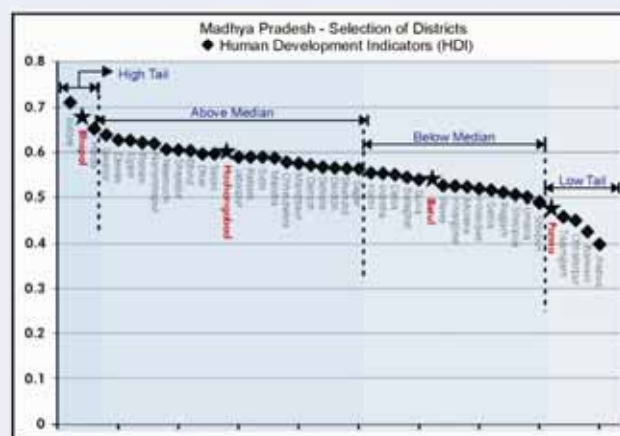
¹⁵ Why could 61% of the students in class 8 not identify the number among the given numbers in which 3 had the greatest place value? Why could students not measure the length of a pencil even at class 8 level? These answers can be obtained through a structured process of speaking to students and carefully listening to their answers. This process can be video-recorded and used in teacher training to help teachers understand how students think and what misconceptions they have. The skill of actually conducting these ‘student interviews’ trains teachers to listen to students - something that usually gets missed in the rush to ‘complete the syllabus’.

States Selected For the Study

S.No	State/UT Name	Population (March 2001 census)	Percentage of total population	Selected/Not Selected	Participated
1	Uttar Pradesh	166,197,921	16%	Selected	No
2	Maharashtra	96,878,627	10%	Selected	Yes
3	Madhya Pradesh	60,348,023	6%	Selected	Yes
4	West Bengal	80,176,197	8%	Selected	No
5	Andhra Pradesh	76,210,007	8%	Selected	Yes
6	Tamil Nadu	62,405,679	6%	Selected	Yes
7	Rajasthan	56,507,188	6%	Selected	Yes
8	Bihar	82,998,509	8%	Selected	Yes
9	Karnataka	52,850,562	5%	Selected	Yes
10	Gujarat	50,671,017	5%	Selected	Yes
11	Chhattisgarh	20,833,803	2%	Selected	Yes
12	Orissa	36,804,660	4%	Selected	Yes
13	Kerala	31,841,374	3%	Selected	Yes
14	Jharkhand	26,945,829	3%	Selected	Yes
15	Assam	26,655,528	3%	Selected	Yes
16	Punjab	24,358,999	2%	Selected	Yes
17	Delhi	13,850,507	1%	Selected	Yes
18	Haryana	21,144,564	2%	Selected	Yes
19	Jammu & Kashmir	10,143,700	1%	Selected	Yes
20	Uttarakhand	8,489,349	1%	Selected	Yes
21	Himachal Pradesh	6,077,900	1%	Selected	No
22	Tripura	3,199,203	0%	Not Selected	
23	Nagaland	1,990,036	0%	Not Selected	
24	Manipur	2,166,788	0%	Not Selected	
25	Meghalaya	2,318,822	0%	Not Selected	
26	Arunachal Pradesh	1,097,968	0%	Not Selected	
27	Mizoram	888,573	0%	Not Selected	
28	Sikkim	540,851	0%	Not Selected	
29	Goa	1,347,668	0%	Not Selected	
		1,025,939,852	100%		

Note: Chandigarh is the capital of the states of Punjab and Haryana, and is a Union Territory, separate from the two states and was included in the study.

Selection of Districts Using Stratification Based on Human Development Index/Literacy Rate - A Sample



A total of 4 districts were selected for Madhya Pradesh, one from each High Tail, Above Median, Below Median and Low Tails of HDI ranking

Chapter 6. SAMPLE DESIGN



6.1 Purpose

Representative and appropriate sample selection and implementation of the same in all states was crucial to achieve the purpose of the project, which was to understand in detail, with respect to specific competencies and in a comparative sense across 18 states, how children in government urban and rural schools are learning in classes 4, 6 and 8. More specifically, the sample design had to be able to:

- Objectively measure the learning levels in the schools across the 18 states.
- Facilitate interstate comparison of student achievement.

6.2 Steps Followed for Finalising the Sample Design

The sample design that was followed is one that is used in large scale educational surveys such as TIMSS. EI's earlier MSB study also followed a similar design. The challenges in SLS for sampling were in deciding the methodology for selection of the districts and in providing representation for the urban and rural schools and appropriate stratification among the primary, upper primary and high schools that had the classes 4, 6 and 8. Educational Initiatives partnered with professors from Indian Statistical Institute, Kolkata to finalise the sampling design. Presentations were also made in National Centre for Educational Statistics (NCES) in Washington, USA for inputs from international sampling experts based on their experience in NAEP and TIMSS. The final sampling plan was also presented to the Ministry of Human Resource, India and suggestions were taken from MHRD, NCERT and EdCIL that enabled finetuning the final sample design.

6.3 Collection of Enrolment Data

Enrolment data from DISE was collected from National University of Educational Planning and Administration (NUEPA). This data was supplemented with the enrolment lists from state governments wherever necessary, for example, all schools in Chandigarh, and class 8 in Surendranagar in Gujarat.

6.4 Methods and Details of Sample Design

Population: Students studying in classes 4, 6 and 8 in government (*DOE and Local Body*) schools in urban and rural areas in the 18 states and 1 union territory participated in the study.

Selection of States: States that have more than 1% of the country's population were selected for participation in the study. The final states tested represented 74% of the Indian population.

Selection of Districts: The Human Development Index (HDI) is an index combining normalized measures of health, literacy, educational attainment, and GDP per capita. It is used as a standard means of measuring human development as per United Nations Development Program (UNDP). In districts where HDI was unavailable, the literacy rate that was published by the census was available. These were recommended by Indian Statistical Institute as a more robust measure to select the sample districts for indicating overall learning in the state than an approach based on geographical-cultural division which was felt to have a higher possibility of misrepresentation of the state's achievement. For e.g., both Hyderabad and Mahbubnagar fall in a Telengana region of Andhra Pradesh and are culturally and geographically alike but very different from each other in terms of development. The criteria that was thus proposed was to select the districts as below:

- Human Development Index/Literacy Rate as the indicator of the district's development to rank the districts and group them into 4 or 2 bins depending on the sample size required per state as High Tail, Above Median, Below Median and Low Tail.
- The final districts for a state were randomly selected from each such group.

However, based on recommendations from GOI, EdCIL and NCERT a composite approach for district selection using HDI/Literacy rate and geographical-cultural representation was attempted to the extent possible. It was also ensured that the final district selection is closer to the proportional distribution of the overall population of the country in the urban and rural areas.

Sampling Plan: It was essential that the sampling plan be kept simple and easy to implement while yielding accurate and efficient samples for both schools and students. The sampling plan proposed a two stage **stratified cluster sampling** design. In stratified sampling the study population is divided into non-overlapping strata and samples are selected independently from each stratum. Selecting whole group of participants i.e. in this case, intact classrooms is known as 'clustering' which serves to increase the statistical efficiency of the sample. Stratification is the grouping of sampling units (for e.g., schools) in the sampling frame according to some attribute of variable prior to drawing the sample (fore.g., urban and rural schools, category of schools as primary, upper primary schools, high schools etc). Stratification ensures adequate representativeness of the sample and improves the efficiency of the sample, thereby making survey results more reliable. Prior to sampling, schools in the sampling frame can be assigned to a predetermined number of stratifications. Combining stratification and clustering approaches allows the study to minimize the overall sample size without compromising its quality.

Sample Size: For a two stage cluster sample, there is no simple equation to calculate sample size as cluster sampling is subject to 'Design Effect' i.e., the variability of sample estimates will be higher than that of simple random sampling, depending on how the clusters differ between themselves, as compared with the within-cluster variation. Different combinations of the number of clusters and number of units selected per cluster will yield different levels of precision. If we assume the variability from the cluster sample to be twice that of a simple random sample, the overall sample size from the simple random sample would need to be doubled to obtain the same level of precision as from a cluster sample. **Based on these, testing covered 1.5 lakh students across the 18 states and 1 union territory to provide score estimates at 95% confidence limit for means $\pm 0.1s$ /Percentages ± 5.0 .**

Sampling Technique: The sample selection method used for this study was systematic *probability-proportional-to-size (PPS)* technique. PPS is a sampling technique in which probability that a sampling unit will be selected in the sample is proportional to some known variable (for e.g. in a population survey, the population size of the sampling unit, in this case, sampling unit is the school). PPS sampling is used:

- when the populations of the sampling units vary, and
- to ensure that every element in the target population has an equal chance of being included in the sample.

This technique, while making the chance unequal for any sampling unit to be chosen, addresses population size. Where the populations of sampling units vary in size and the sampling units are selected with equal probability, the likelihood of elements from a sampling unit with a large population being selected for the survey is less than the likelihood of elements from a sampling unit with a small population being selected. Reducing standard error and bias are two reasons to use PPS. In addition, although weighting can also reduce standard error and bias, PPS sampling is self-weighting and, therefore, simplifies calculations. So another reason to use PPS sampling was to avoid weighting.

Subsequent to stratification and selection of schools, the school enrolment was verified by field personnel prior to testing for participation; else replacement schools selected by PPS were tested appropriately.

Chapter 7. TEST DEVELOPMENT



This chapter describes the process and the steps that formed the basis for the test design and the test development process followed in the Students Learning Study.

7.1 Built on the Earlier Municipal School Benchmarking (MSB) Study

In 2007, Educational Initiatives carried out the MSB study to assess the learning levels of students studying in local body schools in 30 urban towns of the states of Andhra Pradesh, Gujarat, Chhattisgarh, Rajasthan and Uttarakhand. This study tested 26400 students in Language, Maths and EVS in classes 2, 4 and 6.

Table 7.1

Similarities/ Differences	Municipal School Benchmarking (MSB) study	Student Learning Study (SLS)
Year	2006-2007	2008-2009
School system	Government schools and similar private schools	Government schools
Area	30 towns in 5 states of India	48 districts in 18 states and 1 union territory
Schools	Urban	Urban and rural
Classes	2, 4 and 6	4, 6 and 8
Subjects	Language, Maths and EVS	Language and Maths
Languages	3 Languages	13 Languages
Test Components	Written, Group Oral, Individual Oral	Written, Group Oral, Individual Oral

The Student Learning Study (SLS) and the MSB study had a number of similarities and it was logical to build on the work already done in MSB in terms of test design and test development. The steps used in Test Design and Test construction in MSB are as follows:

1. A detailed textbook analysis of the participating states was done to find out: *what the student is expected to know and could do by the end of classes 1, 2, 3, 4, 5, 6 (that she/he could not before that class) and the common minimum curriculum that is followed in each state, in other words, the difference in the curricula across these states class-wise.*
2. The National Curriculum Framework, the focus group documents, the minimum levels of learning (MLLs) and existing research on student learning and pedagogy in India were systematically studied.
3. National level workshops were carried out with subject experts and assessment experts to finalise the competencies and development of items.
4. Some changes to the main tests' design were also done based on a detailed workshop done with experts from different educational organisations such as Vidya Bhawan, Digantar, Homi Bhabha Centre for Science Education, Eklavya, National Institute for Advanced Study and experienced retired experts from NCERT. Feedback was taken on the papers and test design from other organisations such as Centre for Learning Resources, Pune and Azim Premji Foundation.
5. 3 sets of papers were developed for each paper and adapted in 3 languages for the pilot tests to provide adequate pool of items for selection in the main tests. Experts from organisations such as Central Institute of Indian Languages (CIIL) guided the translation, Adaptation and Harmonisation of the versions across the languages.
6. Pilots were done in 3 states and feedback was taken from school teachers. The papers were also analysed for test and item characteristics and fine tuned further.
7. The tests were finally carried out in 30 towns assessing 24600 students.

This rich past experience of the earlier MSB study guided the test development in SLS study in a systematic and successful manner.

7.2 Textbook Analysis

The textbooks of any state are the most important (and often the only) teaching and learning material in schools in a class for a subject. They are a good source of information regarding the curriculum that the children are expected to learn in that academic year. Analyses of textbooks carried out during MSB and other projects by EI helped to map the competencies and skills underlying the curriculum and textbooks of the various State Boards. The test developers studied the analysis to understand what children knew in each class in every state and designed questions appropriately. The textbook analyses also serve as a source on appropriateness of test items and words on cultural and linguistic contexts. The process of analysis entailed:

1. Reading through each chapter/lesson in the textbook
2. Listing down the skills and competencies in every chapter in the textbook
3. Listing down the types of exercises/questions at the end of each chapter/lesson
4. Documenting information such as the degree of match between the MLLs, if stated in the textbook and the content and exercises of the textbook
5. Developing a comparison matrix to identify common content and skill areas between states
6. Creating a comprehensive list of skills and competencies for each class
7. Noting disparities in the curriculum that will need to be known for developing the papers for different states. For example, numerals followed by different states in different classes determine the numerals that will be used in the different version of papers. The table 17 below gives what form of numerals are used by different states in different classes.

Table 7.2

S. No	State	Language	Maths Numerals used in Different Classes in the Participating States							
			Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
1	Andhra Pradesh	Telugu	English	English	English	English	English	English	English	English
2	Assam	Assamese	Assamese	Assamese	Assamese	Assamese	English	English	English	English
3	Bihar	Hindi	English	English	English	English	English	English	English	English
4	Chhattisgarh	Hindi	English	English	English	English	English	English	English	English
5	Delhi	Hindi	English	English	English	English	English	English	English	English
6	Gujarat	Gujarati	Gujarati	Gujarati	Gujarati	English	English	English	English	English
7	Haryana	Hindi	English	English	English	English	English	English	English	English
8	Himachal Pradesh	Hindi	English	English	English	English	English	English	English	English
9	Jammu & Kashmir	English	English	English	English	English	English	English	English	English
10	Jharkhand	Hindi	English	English	English	English	English	English	English	English
11	Karnataka	Kannada	English	English	English	English	English	English	English	English
12	Kerala	Malayalam	English	English	English	English	English	English	English	English
13	Madhya Pradesh	Hindi	English	English	English	English	English	English	English	English
14	Maharashtra	Marathi	Marathi	Marathi	Marathi	English	English	English	English	English
15	Orissa	Oriya	Oriya	Oriya	Oriya	Oriya	Oriya	English	English	English
16	Punjab	Punjabi	English	English	English	English	English	English	English	English
17	Rajasthan	Hindi	English	English	English	English	English	English	English	English
18	Tamil Nadu	Tamil	English	English	English	English	English	English	English	English
19	Uttar Pradesh	Hindi	Hindi	Hindi	Hindi	Hindi	Hindi	Hindi	Hindi	Hindi
20	Uttarakhand	Hindi	English	English	English	English	English	English	English	English
21	West Bengal	Bengali	Bengali	Bengali	Bengali	Bengali	English	English	English	English

7.3 Basis for Test Design

As mentioned before the basis for test-design was competency framework which was designed after analysing the following

- The Minimum Levels of Learning (MLLs).
- The competencies and skills underlying the curriculum and textbooks prescribed by the government of the participating states.
- The learning objectives as stated by the NCERT National Focus Group (2006) position papers in language and maths. Other important skills and competencies identified, based on international best practices.
- Understanding competencies and formats used in other tests across India.

India does not have clearly defined learning standards and as a country the most common curriculum India has are 'Minimum Levels of Learning' (MLL's). However, a criticism that is also levelled against the MLLs is that they tend to be more mechanistic in nature, as they mention in great specificity competencies that must be learnt, while often not mentioning at all the underlying understanding desired. In this context, mapping the curriculum of the states from the state prescribed textbooks assumed greater importance in order to make the tests valid across the different participating states. The National Focus Group position papers released by NCERT in 2006 had explicitly mentioned broad learning outcomes as desired for India in each academic subject which was also used as a guideline while framing the competency matrix for the study. The EI team also reviewed the curriculum of several countries (not as a part of this project) and curriculum frameworks of other international studies such as Third International Mathematics and Science Survey (TIMSS), the Programme for International Student Assessment (PISA), the Progress in International Reading Literacy (PIRLS), etc. Based on all these above, a skill framework was prepared and the specifications for the tests were based on these frameworks. The subject content formed one dimension of the specifications and skills / competencies the other. Details of the test paper design is given in chapter 1.5.

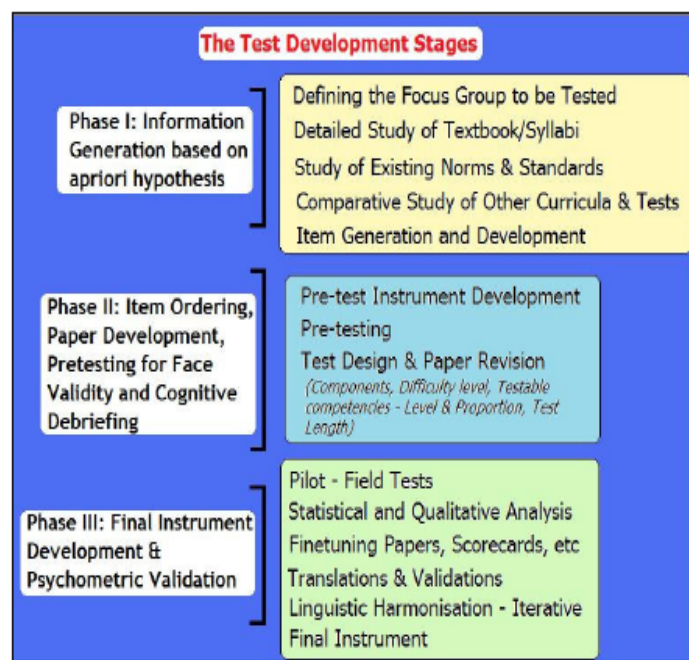
7.4 Process of Test Development

The test development process followed an internationally accepted methodology for finalising the test instruments. This included broadly 3 stages:

Phase I consisted of generating necessary information that will form the groundwork for developing the tests. This included the analysis of textbooks and syllabi, understanding other tests and surveys on student learning and developing the test items.

Phase II consisted of pretesting the developed items to understand the functioning of how well each item traps student understanding of the concept. During pre-tests, cognitive debriefing was achieved through face-to-face interviews with students and teachers for obtaining feedback on the papers. Based on all these inputs the test design and blueprint was fine-tuned.

Phase III included the pilot field tests (given below in detail) that were done on a large scale in 3 states. During the pilot field tests, the tests were carried out in 'as similar' a condition in which the main tests were planned to be carried out.



The papers, scorecards, evaluator manuals, evaluator trainings, standardised test administrations were carried out to ensure appropriate field conditions. The results were further statistically analysed on several psychometric measures. These were done to fine tune the test papers. At each of the stages – during pre-tests, pilot field tests and final paper development, translations, validations and harmonisation of the papers cross languages were done to ensure the comparability of papers. The instruments finalised thus were then used for final testing.

7.4.1 Developing Equivalent Language Versions of All the Tools: The test adaptations and translations follow a rigorous process for compliance with the international test commission (ITC) guidelines. The maths papers were developed in English and are translated into the regional languages based on the guidelines specified for translation which ensures that the original meaning of the question, the reading level of the text, the difficulty level of the item does not change and the likelihood of another possible correct answer for the test item does not arise. Reliability was ensured by further checking the translated papers with language experts and reverse translation into English and synchronisation. This process was iterated for a few times till the standardisation was achieved. An outline of the language test papers was developed in English. Based on this outline, language teams developed the papers in Hindi, Gujarati and Telugu. Subsequently, the Hindi version has been treated as the master and iterative changes are being made in all papers for harmonisation and comparability. A number of critical issues related both to translation validity and appropriateness of certain questions came up at the time of translation, hence the development of papers in multiple regional languages was an iterative process. This process also ensured that cultural adaptations were documented and the test items were acceptable to all of the dialects of the language in which the test is to be administered. Usually the translators were permitted to adapt the text as necessary to make unfamiliar contextual terms culturally appropriate. These rigorous procedures for translation, cultural adaptations, and verification at every stage provided for comparable test items across each participating state. The team involved in this stage of test development comprised Language experts from organizations such as the Central Institute of Indian Languages (CIIL) and translators from the different states. Sample assignments were given to translators and based on the quality of the completed assignments; the selection was carried out to put together the multiple language teams who will work on translation and adaptation of the tool in the different languages.

7.4.2 Pilot Testing and Item Analysis: The purpose of the pilot was to gather information on the functioning of the new items in the tools, feedback for the test administration process and learnings for the field team in aspects related to standardisation of testing procedures. Pilot tests were conducted in 3 languages (Hindi, Telugu and Gujarati). The districts of Medak in Andhra Pradesh, Vadodara in Gujarat and Ghaziabad in Uttar Pradesh were selected for the pilot and permissions were sought from the government departments at the district level. 16 schools (6 urban and 10 rural) were selected in each district with an average class size of 40 students each in classes 4, 6 and 8. The evaluators were college students recruited locally from in each district and trained for conducting the tests. The tests were conducted successfully from the second to fourth week of September 2008 in all the districts. The papers were finetuned and the final test items for the main study were selected on the basis of analysis of pilot test results.

Chapter 8. TEST ADMINISTRATION



Any large scale assessment requires a systematic approach in all its stages i.e., sampling, test design, test development, test administration, test evaluation, data entry, analysis and dissemination. The rigour with which each of these stages is carried out ensures the success of an assessment program. The test administration in the student learning study was a complex exercise. In order to administer the tests across states, a number of processes such as getting permissions from the different state governments, collection of enrolment data, identification of physical location of schools to be tested, verification of attendance data from teacher records, etc were needed to be carried out prior to final test planning, evaluator recruitment, training and testing logistics. This chapter describes in detail the steps involved in implementing and carrying out the study in a standardised manner across all participating states.

8.1 Permissions

As the study was targeting assessment of students in the schools run by the respective State Governments, communications regarding the study was sent followed by presentations to the senior officials of the different State Governments at the level of the Principal Secretary, Education Secretary, Director of Schools and in some cases State Project Directors of SSA for obtaining permissions. Once the state level permissions were obtained, permissions had to be obtained at the district level too for carrying out testing in the primary, upper primary and high schools that fell under their jurisdiction. This was done by meeting and presenting the study individually to the District Education Officers (DEO) of the 48 districts in the participating states. The states of West Bengal, Himachal Pradesh and Uttar Pradesh refused to participate in the study. The study finally tested in 18 states and 1 union territory of India. Permission letters are enclosed in Appendix K.

8.2 Collection of Available Enrolment Data and Statistics in All States

The project team approached the National University of Educational Planning and Administration (NUEPA) for access to raw enrolment data for sampling. NUEPA had subsequently given access to the DISE raw data that has been collected from the various states. This data was checked for the relevant variables and information required for carrying out the sampling plan. Any inadequacy of information observed in this data was supplemented by enrolment lists from the States, wherever necessary.

8.3 Recruitment

Recruitment of Key Persons: A Project Manager (PM) and 5 Zonal Managers (ZM) were recruited to handle the overall coordination for rolling out the SLS tests in the field. For the ease of all field logistics, the participating states identified for testing were grouped into 5 zones and each zone was managed by one ZM.

Recruitment of State Coordinators (SC) and District Coordinators (DC): 21 State coordinators and 60 district coordinators were recruited, one each for the state and district targeted by the study. The pool for hiring was contacted through advertisements, personal contacts and by writing to the management, social work and B.Ed colleges in the district. The applications received were selected and then screened with telephonic interviews and assignments given on planning tasks. The candidates who cleared these stages were finalised after face to face interviews with the Zonal Managers for their respective States and Districts.

Recruitment of Evaluators: The ZMs along with the respective SCs and DCs carried out the recruitment of evaluators. For each of the 48 districts, 20 to 30 evaluators were recruited for the duration of 1-2 weeks. Evaluators were recruited from colleges of education, social work and other social sciences (anthropology etc). Posters were put up in colleges to mobilize

students, and presentations were made to students about the details of the study and how they would benefit by becoming evaluators. A recruitment test¹⁶ was administered, and evaluators were selected based on their:

- performance in the test
- voice assessment (asking them to read aloud a passage in the regional language of the respective state, during which they were graded for clarity, pronunciation, intonation, fluency and adequate loudness of voice)
- zeal, high patience levels and ability to work for long hours.

The selected evaluators were also given a certificate for participation (at the end of the study) apart from a stipend. A copy of the evaluator certificate is enclosed in Appendix T.

8.4 Master Training Workshops

‘Master trainers’ (MT) were personnel who were overall responsible for training the evaluators and ensuring the quality of test administration and evaluation. Master trainers were selected on the basis of the following criteria:

- A basic understanding of the test development process.
- A basic understanding about the need for the adherence of standardised procedures in field during data collection in research studies.
- Fluency and high ability in the regional language.
- Working knowledge of English.
- Ability to inspire and motivate evaluators.
- Willingness to travel to assigned districts for training.

Master trainers were in-house members from the test development and research teams supplemented by expert teachers, freelance language experts from colleges and other research organizations.

Training the Master Trainers (TOMT) was done through two workshops in Ahmedabad and Hyderabad for the duration of three days each. For each state a minimum of 2 to 4 master trainers were identified for the training. During the master trainer workshops, the master trainers, the zonal managers and the state coordinators were thoroughly given exposure to all aspects of the study, test design and test development. They also were trained on the ‘what’ and ‘how to’ aspects of the evaluator training. The master trainers also carried out a field trial test in 3 schools in both Ahmedabad and Hyderabad. They then evaluated the answer responses by assigning the appropriate answer codes as per the score card. This gave them practical insights on the various issues evaluators were likely to face during administration of the tests and evaluation of student responses. The Zonal Managers and State Coordinators were also a part of the master trainer workshop and received training on the do’s and don’ts the evaluators have to follow while visiting the schools during the testing. (*for example, whether evaluators could take food offered by the school, how they should respond if asked for any comment on the practices followed in school, etc*). The master trainers were also provided with a detailed step by step *master trainer manual* on how to do the evaluator training in their respective districts for easy reference. This is enclosed in Appendix M.

8.5 Evaluator Workshops

Standardised test administration was not only important but also imperative for ethical and technical reasons. If performance on the test was influenced by anything other than the attribute being measured then clearly this would reduce the accuracy of the results, which in turn would make the relevance of the results to the assessment an issue. For example, if the test item was measuring the ability to write with correct spellings (dictation) and the item was orally repeated for the student more than a pre-decided number of times, or the administrator mispronounced the item, the variation in item scores as a result of this would invalidate any conclusions on learning that has been arrived at with this item. In this context, evaluator training assumed immense importance both for test administration and evaluation.

¹⁶ A recruitment test was developed (in English and then adapted into regional languages) to check the basic ability of evaluators in Language and Maths as well as the ability to understand the scoring rubric. The candidates were asked to assign answer codes appropriately for actual samples of student work that was given as a part of the recruitment test. The recruitment test and its scoring sheet are enclosed in Appendix P.

Each evaluator was trained for a period of two days in each of the 48 districts. The evaluators were trained by MTs with the support of the ZM, SC and DC. MTs oriented the evaluators about the project and its purpose, on why the training program was necessary, and gave the evaluators a thorough understanding of each paper and its associated scorecard. The evaluators were also given training on how to pace the 'group oral' and 'written' parts of the test, how to read passages and instructions in the paper, how to handle queries from students during the test and how to enter the answer codes for the student responses in the topsheet (data collection sheet with valid answer codes for each question). MTs also clarified any doubts evaluators had on the test design and development or on what the test item was trying to test.

Evaluators were also given training on selecting students for the 'individual oral' test and how each item will be administered and answers coded. A mock test was also organised to assess the quality of each evaluator. Sample student responses were provided for each item from the pre-tested papers and evaluators were asked to assign codes. Training was also given on do's and don'ts every evaluator needed to follow in the school, rapport building activities, plan for classroom seating, carrying and distribution of test booklets, etc. They were also trained on aspects related to arrangements for the test.



The training folders were taken back from the evaluators on the end of day 2 to prevent any leakage of the papers. Feedback was collected in writing from all participants on both days and any gaps they had in their understanding of the process clarified. MTs also gave recommendations to the ZM and SC on which evaluator was suitable for assigning to which class, and who could be made the team leader, etc. based on their qualitative assessment of the evaluators during the training. Some evaluators who did not comply with the requisite rigour were dropped from the programme at this stage.

Between the training and testing most districts had a time gap of 15-20 days and hence stage 2 training was also organised 3 days prior to actual roll out of main tests in each district. The focus of stage 2 training was to refresh evaluator memory on what they learnt in training, in assigning roles, responsibilities for each evaluator, formation of teams for each school and disseminating the testing plan. This training stage was led by the DC. The evaluators were also familiarised with the evaluator code, EI School code and entry of Topsheets I and II. After a gap of one day that was given for evaluators to practice and familiarise themselves with the papers and reading of passages on their own, the testing was rolled out in most of the states on the 3rd day after the stage 2 training. The detailed step by step *evaluator manual* that covered all aspects of administering the tests was used by evaluators as an easy reference during the testing program. The evaluator manual is enclosed in Appendix N.

8.6 Standardisation Processes for Field Operations

School Verification: Before the test was actually rolled out in the schools it was very important to check the school selected for the test. Understanding the physical location of the school and the travel modes, school timings, enrolment of students in the classes to be tested, etc for an idea of the classrooms would later help the DC to plan the appropriate seating of students during testing. Any other issues that may prevent the school from participating in the research study could also be brought to the fore. The SC and DC (in a few cases, designated team leaders) visited every sampled school prior to testing and informed them about the project, took permission from the school authorities and collected requisite information along with the Head Teacher's contact details.

Test Administration: The tests were conducted in two phases. In the first phase it was conducted in 12 states and Chandigarh from January–March 09 and in the second phase it was conducted in 6 states from July–September 09. During testing:



- One evaluator was assigned for administering the test for each classroom.
- Immediately after the school assembly, evaluators entered the classes and started rapport building activities with students. After a non-threatening comfort zone was established with students, evaluators started the tests.
- In each class, the evaluator administered both the papers on the same day with a minimum gap of 30 minutes between the two papers. The language test was administered first followed by the maths test and the student questionnaire.

- The students were provided with the necessary writing implements like pencils and erasers before the start of each test.
- Oral tests in language were held after the testing for class 4. From each classroom 8 students were identified by skip counting in multiples of 5 or 3 as appropriate from the attendance list for this. While administering the test to each student, the evaluator immediately marked the relevant answer code for the student for each item that was administered in TOPSHEET II in the appropriate place.



As far as possible, the DCs visited all the testing locations every day for smooth running of the testing. To ensure the smooth administration of the tests and as a quality control measure, testing sites were also visited by ZMs on the first day of testing in each state.



The evaluators were given answer sheets in the morning and they were collected back in the evening. It was important to ensure that no question paper remained on the premises or was copied as the test dates were different across schools. At the end of the day, all the question papers and answer sheets were collected and counted. No evaluator was allowed to carry the testing material to her/his home. The DC used detailed checklists and maintained track of the number of test papers, answer sheets, etc. that were used by each school. The PM who managed the testing and resources centrally from Hyderabad was updated at the end of each day. To achieve maximum transparency on the process, an online system

was also developed for tracking the daily reporting from the field. Each day, the DCs entered on a daily basis the total enrolment, number of students present, number of students tested and test papers used for that day to enable tracking of testing across the country.

Answer Coding: Test papers for every class were supplemented with a **Score Card**. The Score Card provided a question-wise rubric (with scores / codes / answers) for how each response was to be scored. This process ensured that evaluators do not “correct” questions and award marks, but only assign codes based on student responses. This minimises evaluator bias in correcting a test paper. The Score Card captured common mistakes that children tend to make while answering questions. The data that was collected by this process was critical for the diagnostic feedback and analysis. This form of scoring is common practise in international tests like the TIMSS. The evaluators were provided with a ‘**topsheet**’ – which was an OMR sheet. The OMR sheet had pre-designed valid codes entered for each question in the paper. The evaluator had to darken the oval containing the appropriate answer code for the answer given by the student. Use of OMR sheet eliminated one layer of error that occurs usually in large testing programs during data entry. A sample of the score card and the topsheet is provided in Appendix W.

Each day of testing was followed by coding the next day for the schools tested the previous day. Coding was done at the district Headquarters in a pre-designated place where evaluators assembled and carried out the coding activity under the supervision of the District Co-ordinator. ZMs also visited the evaluation venues on the first day of coding in each state and carried out spot checks on topsheet entries made by evaluators for specific student responses. The doubts during evaluation were cleared by DC/SC or by consulting the MTs and the core team working on the project. The same evaluator administered all the tests to the students of a particular class. On the first day of coding, evaluator errors were calculated by the DC and was discussed with the ZM. Evaluators were sensitized to wrong coding (if any) on the first day itself, and asked to make the corrections so that it did not get repeated subsequently.

Background Questionnaire Administration: Background Questionnaire or the “School Information Schedule” collected information on the school resources, head teacher details and teacher information. The DCs or team leaders collected the information from schools.

Printing and Despatch: The test instruments were centrally printed in Hyderabad and Ahmedabad and despatched to the different districts. This enabled standardising the aesthetic aspects related to the paper used and print quality across the different language versions of the papers.

Manuals and Logistics Tracking Forms Used: The mammoth operation of testing was made smooth and standardised by use of various manuals which laid the guidelines critical to all aspects of training and testing. The logistics tracking forms enabled efficient processes for roll out of the tests across the different states. These are listed below:

Table 8.1

S.No	Manuals	Appendix
1	Master Trainers Manual	M
2	District Coordinators Manual	O
3	Evaluators Manual	N
4	Frequently Asked Questions About The Project	L

Table 8.2

S.No	Forms Used	Appendix
1	Stock Received Verification Form	O
2	Testing Dates Chart	O
3	Class and Role Allotted	O
4	Evaluator Label and Testing Table	O
5	Evaluator Details and Attendance Sheet	O
6	DC Daily Report form	O
7	The Enrolment Number Feedback Form and FAQ	S
8	Evaluator Training Attendance sheet	R
9	Student Attendance Sheet	O
10	Evaluator Training workshop feedback form	Q
11	Master Trainer's feedback form	Q

Chapter 9. REPORTS AND ANALYSIS



9.1 Methods of Data Handling

The evaluators marked the codes for the responses given by the student in pre-designed OMR sheets that listed the valid codes for each question. The OMR sheets were scanned and the data taken into the database. This was further cleaned using special data cleaning programs that identified data errors such as duplicate/triplicate records, errors in fields such as the school code, district and state information, class, language and maths test attendance, etc. The errors reflected in the error instances file were rectified in a systematic manner through physical checking of the OMR sheets and supported by other records such as attendance sheets, district coordinators' daily reporting system and the actual answer scripts in required cases. The final test data that was scored and used for further analysis in all the states has 0% error. The background questionnaire data was punched in through data entry operators and checked for errors.

9.2 Analysis Methods

After scoring of the data, a number of reports were generated to check for the test characteristics – number of observations, mean score and standard deviation (state-wise), reliability, etc. A range of diagnostic statistics was also derived to evaluate the psychometric characteristics of each test item. This enabled the detection of any uncharacteristic behaviour of an item such as extreme difficulty or vice versa, very low discrimination power, as well as any issue that may have been possibly caused due to inaccuracies in translated versions, printing, etc.

The data was then scaled using Item Response Theory (IRT). IRT model is a 'latent variable' model that describes the probability that a student will respond in a specific way to an item in terms of the respondent's proficiency, which is an unobserved or 'latent' trait and various characteristics (or 'parameters') of the item i.e., IRT puts item and students on the same scale. Thus the difficulty of an item and the ability of a person can be meaningfully compared. The latent variable, expressed as θ , is a continuous one-dimensional construct that explains the covariance among item responses. Students at higher levels of θ have a higher probability of responding correctly or endorsing an item. The student scores were scaled to a mean score of 500 and SD of 100 as per internationally acceptable conventions. The IRT item parameters and the matrix plot for each test are enclosed in Appendix G and H respectively. Scale Anchoring of questions to understand what students know and do at different ability levels were carried out as described in chapter 4.

The background questionnaire data and the test data also were subjected to a number of statistical analyses such as t-tests, Anova, Tukey's HSD, Correlation, Multiple Linear Regression, etc in order to find the relationships that exist in the observed data.

Distracter analysis carried out at item and option levels enabled the identification of misconception and common errors.

A small sample of student responses on the free response writing questions in language were also analysed qualitatively to understand how students perform in these items. The details of this analysis and the findings are given in chapter 2.4.

BIBLIOGRAPHY



1. Allington, L. R., & Johnston, H.P. (2000), "What Do We Know about Effective Fourth-grade Teachers and their Classrooms?"
2. Adams, R.J., Wu, M.L., & Macaskill, G. (1997), "Scaling Methodology and Procedures for the Mathematics and Science Scales" in M.O. Martin and D. L. Kelly (Eds.), *TIMSS Technical Report Volume II: Implementation and Analysis*. Chestnut Hill, MA: Boston College.
3. Adams, R.J., & Khoo, S. (1993), "Quest: The Interactive Test Analysis System" Melbourne: Australian Council for Education Research.
4. Balanskat, Blamire, Kefala (2006), "The ICT Impact Report".
5. Barber, Mourshed, Whelan (2007), "Improving Education In The Gulf".
6. Beaton, A.E., & Johnson, E.G. (1992), "Overview of the Scaling Methodology used in the National Assessment". *Journal of Education Measurement*.
7. Cheng (2004), "New Paradigm for Re-engineering Education".
8. Currie, T. (1998), "Early Test Scores, Socioeconomic Status and Future Outcomes".
9. Educational Initiatives (2007), "A Benchmarking Study of Student Achievement in Local Body Schools of Large Towns of Select States".
10. Educational Initiatives (2006), "Student learning in Metros".
11. Educational Initiatives (2009), "Bhutan's Annual Status of Student Learning".
12. Ellis (2006), "The Training and Development Agency for Schools".
13. Elmore (2004), "School Reforms From The Inside Out".
14. Eurydice (2005), "Key Data on Education in Europe".
15. Hanushek E. A., & Woesmann L. (2007), "The Role of Education Quality for Economic Growth", Policy Research Working Paper Series 4122, The World Bank.
16. Johnes & Johnes (eds) (2004), "International Handbook on the Economics of Education".
17. Karthik M., and Michael K., (2006), "Public and Private Schools in Rural India".
18. Lord F.M. (1980), "Application of Item Response Theory to Practical Testing Problems". Hillsdale, NJ: Erlbaum Associates.
19. Lewin K.M. (1999), "Counting the Cost of Teacher Education Cost and Quality Issues", Center of International Education, University of Sussex, Brighton, MUSTER Discussion Paper I.
20. McKinsey Report (2008), "How the World's Best Performing School Systems Come out on Top".
21. Mullis I. V. S., Martin M. O., Kennedy A. M. & Pierre F. (2007). "PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries." TIMSS & PIRLS, International Study Center, Chestnut Hill, MA: Boston College.
22. National Centre for Education Statistics (2006), "A Closer Look at Charter Schools Using Hierarchical Linear Modeling".
23. National Centre for Education Statistics (2003), "A Closer Look At Charter Schools: Results From The NAEP Pilot Study".
24. National Commission on Excellence in Education (1983), "A Nation at Risk: The Imperative for Educational Reform: US Department of Education".
25. National Council on Teacher Quality (2004), "Increasing the Odds: How Good Policies Can Yield Better Teachers".
26. NFER (1997). "Trends in Standards in Literacy and Numeracy in the United Kingdom".
27. Phillips (2007). "Linking NAEP Achievement Levels to TIMSS".
28. PIRLS (2001), "International Report", Chestnut Hill, USA.
29. Pratham (2005-2008), "Annual Status of Education Report".
30. Schleicher (2006), "Why Education is Key for Europe's Success".
31. Scientific American (2001), "Does Class Size Matter".
32. Schmidt, W. & Cogan, L. (1996), Development of the TIMSS Context Questionnaires, in M.O. Martin & D.L.Kelly (Eds.), *Third International Mathematics and Science Study Technical report, Volume I*. Chestnut, MA: Boston College.
33. The World Bank (2004), "Books, Buildings, and Learning Outcomes: An Impact Evaluation of World Bank Support to Basic Education in Ghana", Report 28779, Operations Evaluation Development.
34. TIMSS (1998a). Survey Operations Manual (Doc. Ref. No. 98-0026). Prepared by the International Study Center at Boston College. Chestnut Hill, MA: Boston College.

35. TIMSS (1997c). School Coordinators Manual (TIMSS 1999 Doc. Ref. 98-0024). Prepared by the International Study Center at Boston College. Chestnut Hill, MA: Boston College.
36. TIMSS (1997d). Test Administrators Manual (TIMSS 1999 Doc. Ref.98-0025). Prepared by the IEA Data Processing Center. Chestnut Hill, MA: Boston College.
37. TIMSS (1997). TIMSS 1999 School Sampling Manual-Version 2 (Doc. Ref.:TIMSS1999 97-0012). Prepared by Pierre Foy, Statistics Canada. Chestnut Hill, MA: Boston College.
38. TIMSS (1998a).Survey Operations Manual-Main Survey (Doc. Ref.:TIMSS1999 98-0026). Prepared by the International Study Center, Chestnut Hill, MA: Boston College.
39. Training and development Agency for schools, Press Release (11 August 2005).
40. UNESCO (2005), "EFA Global Monitoring Report".
41. UNESCO (2004), "Education for All. The Quality Imperative."
42. UNICEF (2000), "Defining Quality in Education", Working Paper Series.
43. Washington School Research Center (2005), "The Power of Early Success".
44. <http://www.arthurhu.com/index/classize.htm>
45. <http://nces.ed.gov/Surveys/PIRLS/released.asp>
46. <http://timss.bc.edu/timss2003i/released.html>
47. <http://timss.bc.edu/timss1995i/items.html>