GUIDELINES FOR LARGE-SCALE LEARNER ASSESSMENTS

Practices for Design, Implementation and Use of Assessments by States

March 2014

CENTRAL SQUARE FOUNDATION

FICCI

GUIDELINES FOR LARGE-SCALE LEARNER ASSESSMENTS

Practices for Design, Implementation and Use of Assessments by States

March 2014

ACKNOWLEDGEMENTS

The FICCI School Education Committee would like to express heartfelt thanks to Vatsala Deora and Anastasia Aguiar (Central Square Foundation) for the authorship of this report. We also deeply thank several individuals for continuously supporting our work – Ajay Batra, Aanchal Chomal and Prachi Gaur (Azim Premji Foundation), Sridhar Rajagopalan and Vyjayanthi Sankar (Educational Initiatives), Jim Tognolini and Isabel Sutcliffe (Pearson), Sandeep Aneja, Gaurav Jain and Raj Shastri (Kaizen PE) and Ratna Dhamija (ACER) for guiding the research and writing. We are extremely thankful for the ongoing support from Shobha Mishra Ghosh and Sunita Mohan (FICCI).

We especially appreciate the support of state government officials who allowed us to observe and understand the mechanisms of their assessment programmes – specifically, the Karnataka State Quality Assessment and Accreditation Council, Madhya Pradesh Pratibha Parv and Gujarat Gunotsav. Your time and patience were crucial to capturing the essence of these incredible initiatives. We would also like to thank Amitav Virmani, Kanupriya Misra and Syed Farooq Hasan (ARK), Ajay Batra and Rishikesh (Azim Premji Foundation) and Vishnuneeth Agnihotri and Rohit Mehta (Educational Initiatives) for organizing these state government interactions.

The authors wish to acknowledge the following colleagues for their contribution to the report – Azad Oommen for his overall inputs and editing support and Radhika Kapoor and Shrutipriya Dalmia for managing the production process with support from Rakesh Satone and Abhishek Batra (Rakesh Press). The authors are thankful to Neha Tulsian for guiding the report design.

CONTENTS

ABBREVIATIONS AND ACRONYMS

ACARA	Australian Curriculum Assessment and Reporting Authority
ACER	Australian Council for Educational Research
APF	Azim Premji Foundation
AP RESt	Andhra Pradesh Randomized Evaluation Study
ASER	Annual Survey of Education Report
CCE	Continuous and Comprehensive Evaluation
CIIL	Central Institute of Indian Languages
DIET	District Institute of Education and Training
DoE	Department of Education
EGRA	Early Grade Reading Assessment
EI	Educational Initiatives
FICCI	Federation of Indian Chambers of Commerce and Industry
IEA	International Association for the Evaluation of Educational Achievement
IRT	Item Response Theory
MLL	Minimum Learning Level
NAPLAN	The National Assessments Program – Literacy and Numeracy
NAS	National Achievement Survey
NCERT	National Council for Educational Research and Training
PIRLS	Progress in International Reading Literacy Study
OMR	Optical Mark Reader
PISA	Programme for International Student Assessment
SCERT	State Council for Educational Research and Training
SLS	Student Learning Study
SMC	School Management Committee
SSA	Sarva Shiksha Abhiyan
TIMSS	Trends in International Mathematics and Science Study

FOREWORD

In the coming decades, India hopes to rank amongst the best global economies while being proud of a society that thrives on values of democracy, justice and equity. With this mandate, we need to ensure that our future generation is armed with the knowledge, skills and attitudes that empower them to be productive individuals and citizens. Our schooling system needs to ensure that our students are learning, are grounded in Indian ethos and are being prepared for the world. Though international and national benchmarks of learning can provide a glimpse of the relative performance of countries and states, the onus lies with each state government to holistically evaluate the level of learning of their schooling system and plan resources, interventions and funds to improve this.

With large-scale assessment, each state can begin the journey to emphasize the provision of a high-quality education for each and every student, by assessing where they are and targeting where they want to reach. A well-designed and administered large-scale assessment can provide regular, system-level information on student learning to gauge overall levels of achievement and the performance of specific sub-groups that may be at higher risk of falling behind. Furthermore, if data from such an assessment is analysed and released in a timely and useful manner, government can determine the effectiveness of its policies and alter them accordingly.

These guidelines aim to support this culture of evidence in building strong learning outcomes in states. With in-depth explanations of the steps involved in assessment design, implementation, analysis and result dissemination, this document can act as a companion to state decision-makers who are encouraging large-scale assessment in their constituencies.

On behalf of the FICCI School Education Committee, we encourage central and state stakeholders to analyse and contextualize these guidelines as they take the lead in bringing evidence-led reform in education for better learning for our children.

Gowri Ishwaran
Co-Chair,
FICCI School Education Committee

Ashish Dhawan
Chair,
FICCI School Education Committee

Prabhat Jain
Co-Chair,
FICCI School Education Committee

Internationally, large-scale assessments have seen significant success as levers to manage policy evaluation and accountability, especially planning resources, goal setting and allocation of funds.

EXECUTIVE SUMMARY

These guidelines showcase best practices for the design, implementation and analysis of state-led large-scale assessments. They are supplemented with case studies of assessments conducted in India and around the globe, to bring forward how decision-makers have designed outcomes-focused solutions in various contexts.

Large-scale learner assessments are a mechanism for states to gauge how well learning is happening in their districts, blocks, schools and classrooms. The 12th Five Year Plan states that there is a need to *"improve learning outcomes that are measured, monitored and reported independently at all levels of school education".* This document presents global and national practices to aid states to vision, design and build administration structures, conduct score analysis and use results of large-scale learner assessments.

The guidelines highlight key reasons for the need for large-scale learner assessments and emphasize the importance of measuring learning outcomes in a valid, reliable and recurring manner.

A state must begin the large-scale assessment design process with a clear purpose in mind. Assessments may be designed to evaluate the system, to hold it accountable or to define strategies for student improvement. This purpose needs to be determined clearly by a steering committee that includes key stakeholders such as the Department of Education (DoE), the State Council for Educational Research and Training (SCERT) and the District Institutes of Education and Training (DIET).

Internationally, large-scale assessments have seen significant success as levers to manage policy evaluation and accountability, resource planning, goal setting and allocation of funds. In India, states have the opportunity to utilize large-scale assessments specifically for this purpose, while focusing on the classroom-based Continuous and Comprehensive Evaluation (CCE) structure for improvement in individual student learning.

Assessment design and administration are crucial to ensuring the validity and reliability of data. The state must decide the class-level, frequency and population of the assessment to guide the creation of the assessment framework. This document then describes the exact content and cognitive domains being tested and forms the basis of test item design, development and field pilots.

The administration of a large-scale assessment is a complex task and involves the participation of many stakeholders. The state may choose to engage external administrators and scorers, or have teachers conduct tests in class independently. In either case, the administrator requires significant training and a standardized manual to ensure the assessment is conducted in a singular manner across the state. Further, as a learner assessment is accompanied by a background questionnaire, to ascertain contextual factors that may determine student performance, the principal and other staff members also need adequate guidance ahead of time.

The data captured from a large-scale assessment can be cumulated in various ways – at the student, teacher, class, school, cluster and block or district level. Trends in the performance of sub-groups such as girl students and schools in rural surroundings and the like, can also be ascertained. The guidelines discuss some common recommendations to ensure that the data is scored and collated reliably. Additionally, some of the most common reporting and result dissemination formats, such as press releases, summary reports, conferences, are also described.

The life cycle of the large-scale assessment process, from design of the instrument to dissemination of results, must be accounted for in the full cost structure and a breakdown of the various cost heads is discussed. The final budget can represent between 0.3% to 2% of a state's education budget.

Finally, the guidelines describe the three key uses of assessment results, i.e. policy recommendations, teaching interventions and public awareness programmes. Each can have a significant impact in improving learning outcomes over time.

Overall, these guidelines recommend that implementing a large scale-assessment is a journey states must embark upon. As a nation, we hope to build a culture of evidence, ensuring that every child is learning. This has been reinforced by the Ministry of Human Resource Development (MHRD) and Sarva Shiksha Abhiyan (SSA), as they have emphasised measures to track learning-outcomes in state planning. This culture will be built on a foundation of a strong, national survey - such as the National Achievement Survey (NAS) conducted by the National Council of Educational Research and Training (NCERT) - one that is administered frequently with a nimble distribution and management strategy. This survey can share results with state and district authorities in a timely and relevant manner and benchmark the performance of states to encourage peer-learning and collaboration.

This foundation of a national survey must be supported by state-level census assessments, that are conducted every one or two years. This will lead to transformational change - rather than incremental change - as these assessments are linked directly to state curricula and context and provide data that can be shared with every district, block, school and community, whilst also influencing state policy and planning. This will be an arduous, intensive task, but it is imperative that states take the first steps to design and administer purpose-led instruments, analyse and and report the data and build the capacity of their officers to learn and continuously improve.

Section One

Assessment Vision

ASSESSMENT VISION

This section examines the need for a state-led large-scale assessment and differentiates it from current examinations that students take. It explains the purpose of an assessment and highlights the structure of the institutional mechanism to design and deploy.

The Need for State Assessment

Many countries have begun to realize that there is a need to create and implement standards for holistic quality in order to make schools vibrant learning organisations. Quality standards should take into account tangible inputs such as infrastructure, enrolment, teacher-pupil ratio while also measuring intangible processes and outcomes such as student learning levels, pedagogic processes and overall school culture. While measuring some of these intangible parameters of quality is complex, there is a need to move towards a 'culture of evidence'.[1]

In India, a growing body of evidence around learning outcomes, i.e. the knowledge and skills that students have acquired as a result of their exposure to schooling[2], indicates

a dire current state. However, there is limited empirical evidence available to support this claim, as a majority of the state governments in the country do not monitor student learning in a periodic, valid and reliable manner.

As the focus of governments shifts from access to the quality of education, strong empirical evidence about student learning is required at every stage in the schooling process. This evidence, along with teacher assessment, school assessment and ongoing reform, can guide data-driven decisions about curriculum, pedagogy, classroom learning and teacher training.

The traditional Indian examination structure does not suffice to track learning outcomes on a systemic level, because:

- The purpose of internal school assessment is to evaluate the achievement of individual students and not the system as a whole.
- The focus, design and difficulty of these examinations varies greatly and does not take into account background factors that may impact learning.
- Common board examinations are conducted only in Class 10 and Class 12, which is the end of a child's schooling career. The results are 'high-stakes' for the students because they determine future course of study or employment. Hence, examinations are designed to allow the maximum number of students to qualify, and not specifically to distinguish between them.
- Students take a different set of question papers each year, with no unifying rubric to allow comparison of scores across years.

A wide variety of other assessment activities have been executed in India. NCERT conducts the NAS, initiated under Sarva Shiksha Abhiyan (SSA) on a four-year cycle. Non-government efforts also include the Annual Status of Education Report (ASER), an ASER Centre led effort in rural India, the Learning Guarantee Program (LGP) managed by Azim Premji Foundation (APF) and multiple urban and rural achievement surveys performed by Educational Initiatives (EI).

Though the results of such assessments all point in the same direction, the approach and content that impact design are pre-determined by the commissioning agency and this may not match the context of a particular state or the priorities of its policymakers.

Therefore, it is essential that states implement their own assessments. These can share the key characteristics of their national and international counterparts but the research questions should be set locally and the assessment instruments developed should closely fit state curricula and intended educational standards.[3]

Kellaghan and Greaney (2001b, 2004) describe that large-scale student assessments would help states understand:

1. Level of learning in the education system (with reference to general expectations, aims of the curriculum, preparation for further learning or preparation for life).

2. Particular strengths and weaknesses in students' knowledge and skills.

3. If particular sub-groups in the population perform poorly and if disparities exist, for example, between the achievements of:

    - Boys and girls
    - Students in urban and rural locations
    - Students from different language or ethnic groups
    - Students in different regions of the state

4. Factors that are associated with student achievement and to what extent achievement varies with characteristics of the learning environment such as school resources, teacher preparation and competence and type of school or with students' home and community circumstances.

5. Compliance with government standards in the provision of resources e.g. textbooks and teacher qualifications.

6. Trends in student learning over time. This may be of particular interest if reforms of the education system are being undertaken. Answering the question requires carrying out assessments that yield comparable data at different points in time.[4]

The Purpose of State Assessment

A large-scale assessment will generate a significant amount of data about student learning levels. The purpose of such an assessment must be clear in order to ensure that the test is appropriately designed and valid evidence is collected.[5] Additionally, if the test is designed for a specific purpose, the results should not be used for a different purpose, as it is likely that any inferences made based on test results will not be accurate or valid for other purposes.[6]

Broadly, there are three primary objectives of a large-scale assessment:

- **Evaluation** – Large-scale assessments are often a major monitoring mechanism for a system. Monitoring and evaluation refers to collecting and analysing data to check performance against goals and to take remedial actions if needed.[7] These 'goals' are national or state learning standards and thus the assessment needs to be aligned very directly with the curriculum. In Australia, the The National Assessments Program – Literacy and Numeracy (NAPLAN) compares student results to minimum learning standards and monitors student progress accordingly. Individual states in India follow specific curricula under the guidance of National Curriculum Framework and hence nodal monitoring points can be at the state level.

- **Accountability** – Where assessment is used to hold any part of the system accountable, there need to be clear consequences of the evaluation. Globally, this has taken various forms. First, there is a growing trend of public reporting, including the publication of standardized student assessment results at the school level for use by parents, government officials, the media and other stakeholders, the publication of school inspection reports, school annual reports and system

level reports providing an assessment of the state of education. Second, assessment results are increasingly used to reward or sanction the performance of individual school agents. This goes alongside the expansion of school external evaluation and teacher appraisal procedures.[8]

- **Improvement –** Countries utilize assessment results for formative purposes, providing feedback to teachers on specific student performance. In this case, the results must be presented to teachers, school leaders and government officials in a meaningful manner, such that they can be readily utilized. In France, results from diagnostic tests are used to form groups of students for whom personalized assistance programmes are offered.[9]

There is significant international discourse about using a single standardized assessment for multiple purposes. Most countries, such as those highlighted in the supporting case studies with these guidelines, have several, sometimes conflicting, objectives. This can increase the ambiguity and validity of results.

In India, it may be difficult to implement strategies for classroom improvement directly from a large-scale assessment. However, the CCE structure ensures that teachers follow periodic formative assessment for specific learning checks. Hence, a large-scale assessment informs and monitors the system and evaluates the effectiveness of policy and resource utilization.

Institutional Mechanism for State Assessment

In building the vision for large-scale assessment, it is crucial for a state to identify which institution or team within the government education structure will be responsible for execution. The subject of assessments is a specialized one, requiring research focus and narrow expertise in areas like item development, Item Response Theory (IRT), computer adaptive testing and test equating. A number of civil society organisations - government agencies, university departments, companies and others have expertise in these subjects.[10]

In India, the NCERT has the central mandate to design and conduct large-scale assessments and achievement surveys. In this capacity, it carries out the NAS across the country. Within states, each SCERT is decreed to manage the process. However, currently many of these institutions may not have the training, resources or bandwidth to accomplish this process.

Many countries have instituted independent agencies to conduct national assessment. The Australian Curriculum Assessment and Reporting Authority (ACARA) is an independent authority providing a rigorous, national approach to education through the national curriculum, national assessment programme and national data collection and reporting programme in Australia. Alternatively, some countries have established centres of excellence for assessment within their existing government structures. In both cases, governments have worked closely with external subject matter experts, especially in the early years, to build capacity and continuously improve.

Bangladesh designed and conducted a national sample assessment of students in Classes 3 and 5 in 2006. The Directorate of Primary Education led this effort and, through a rigorous tender process, engaged the Australian Council for Educational Research (ACER) to

provide advice and quality assurance. ACER's consultancy involved providing short-term practical support in developing test items, conducting workshops with the curriculum and design teams, defining the test administration guidelines and preparing the surveyor manual. Furthermore, ACER identified long-term strategies to support capacity building within the system and gradually transferred full responsibility to the national team. Similar assistance is also required in the result analysis, reporting and dissemination process.

In Puducherry, the SSA and the DoE have brought together a group of 45 people, comprising subject teachers, DIET faculty and curriculum developers, to form a core team to lead large-scale assessment. The DoE recognizes that they require training in test development, administration and analysis and have collaborated with APF to provide continuous support.

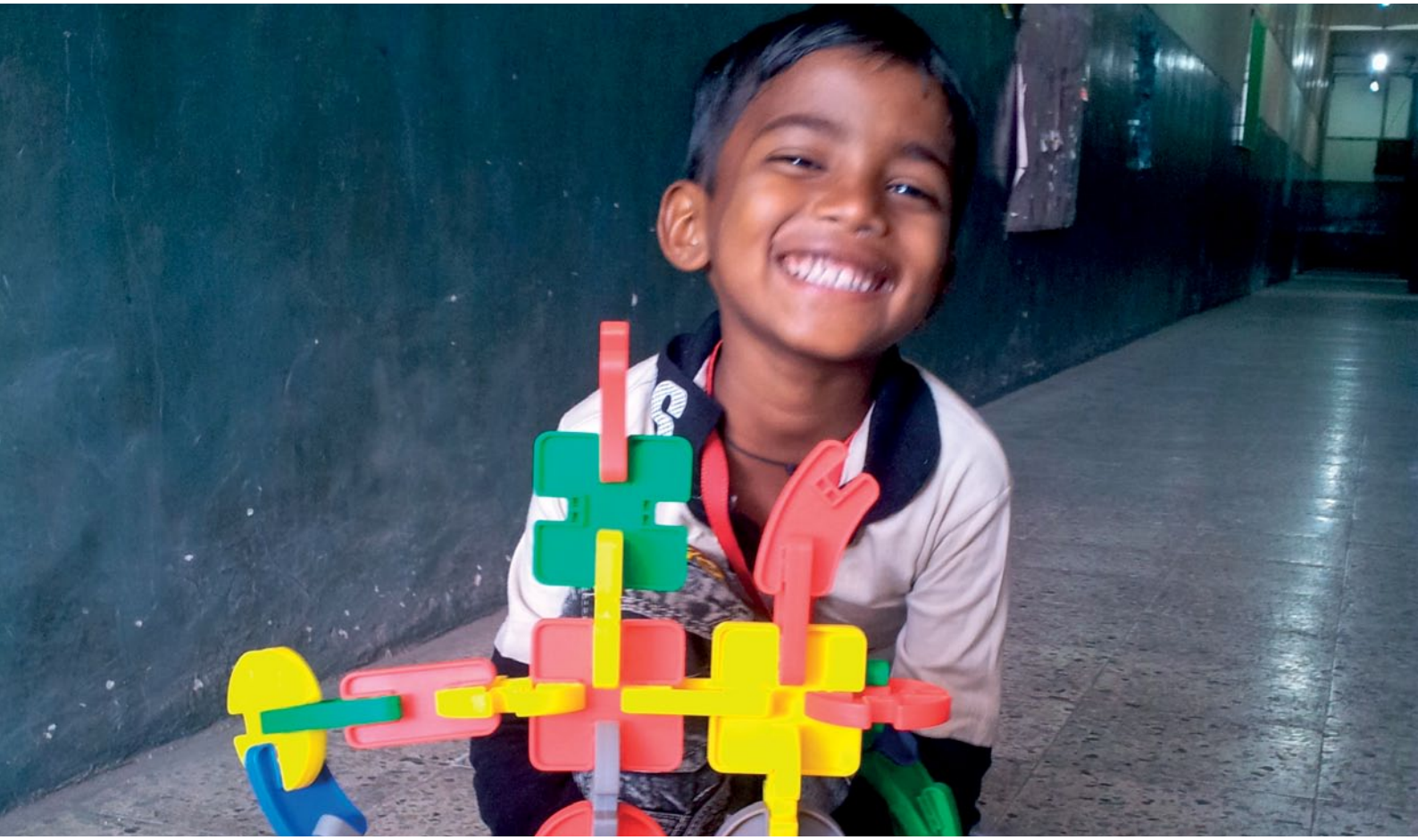
The state government should also appoint a steering committee to provide guidance to the assessment agency. The committee should ensure that the assessment has priority status and addresses key policy questions. It could also help resolve serious administrative and financial problems that might arise during the implementation of the assessment. Giving this committee a degree of ownership over the direction and intent of the assessment also increases the likelihood that the results of the assessment will play a role in future policy making.[11]

The composition of a steering committee will vary from state to state, depending on the structure within the education system, but could include:

- The State Principal Secretary
- The SCERT Director
- DIET Principals and Faculty
- Representatives from the Central Government
- Teachers
- Representatives from the Teacher Union bodies
- School Management Committee members
- Representatives from Teacher Training Institutes
- Civil Society Organisations
- Assessment Partner Organisations

Section Two

Assessment Design



ASSESSMENT DESIGN

This section delves into the specifics of assessment design. It defines and discusses the key elements of design, including the population, frequency and assessment framework. It also examines best practices in item development and background questionnaire planning.

Principles for State Assessment

A state should approach a large-scale assessment with some key questions:

- **Why do we need to conduct this assessment?** This refers to the purpose of the assessment, as discussed in the previous section.
- **Who will take this assessment?** This signifies the population and class-level at which the assessment will be conducted. This is discussed in the `Elements of Assessment Design' sub-section.
- **When will the assessment take place?** This refers to the frequency and timing of the assessment. This is also discussed in the `Elements of Assessment Design' sub-section.

- **What will be assessed?** This will become the assessment framework, i.e. the content and cognitive domains that are tested. This is elaborated further in the 'Assessment Frameworks' sub-section.
- **How will the assessment findings be used?** Before a state conducts an assessment, it must have a plan for how it will analyze, report and disseminate findings and how these will be used to guide policy changes or learning improvement interventions. This is further explained in the 'Assessment Reporting and Use of Results' section.

Additionally, there are some overarching principles to large-scale assessment design that must be kept in mind. These include:

- Assessment should be 'low-stakes', i.e. they do not result in any rewards or consequences for the student.
- There should be no fear of consequences in the minds of teachers or school administrators.
- The state should have a long-term commitment to conducting the assessment in a periodic manner and using results in planning resources and monitoring the success of its policies and interventions.
- The assessment should be conducted with participation and buy-in from all relevant stakeholders, including teachers, principals, civil society, policymakers, teacher trainers and School Management Committee (SMC) members.
- The state should conduct the process in such a manner that it gains expertise in conducting the assessment from external providers who are initially commissioned to assist in the design, administration and analysis of the process.
- The test development process should ensure there are adequate items to accurately assess students across knowledge, inference and critical reasoning domains. Additionally, these items should be comparable across assessment cycles.
- Findings should be communicated with stakeholders in a timely and useful manner, with provision in the assessment budget for reporting and dissemination of results.

Elements of Assessment Design

The main elements of assessment design are summarized below (Table 1):

Table 1:
Key Elements of Assessment Design

Section	Elements
Purpose	The DoE, in collaboration with the steering committee and key stakeholders, establishes the purpose of the assessment
	The DoE appoints and provides funding to either an implementing agency within the ministry or an independent external body (such as, a university department or research organisation)
	The DoE determines the policy needs to be addressed in the assessment, sometimes in consultation with key education stakeholders (including teacher representatives, curriculum specialists and parents)
Principles	The DoE, or a steering committee nominated by it, identifies the population to be assessed (e.g. Class 4 students)
	The DoE determines the area of achievement to be assessed (e.g. literacy, numeracy)
	The implementing agency defines the area of achievement and describes it in terms of content and cognitive skills
Design	The implementing agency prepares achievement tests, supporting questionnaires and administration manuals and takes steps to ensure their validity
	The tests and supporting documents are pilot-tested by the implementing agency, steering committee members and other competent bodies to: a. Determine curriculum appropriateness b. Ensure that items reflect gender, ethnic and cultural sensitivities
	The implementing agency selects the targeted sample or population of schools or students, arranges for printing of materials and establishes communication with selected schools
Administration	The implementing agency trains test administrators
	The survey instruments (tests and questionnaires) are administered in schools on a specified date under the overall direction of the implementing agency
	The implementing agency collects survey instruments, scores them and and prepares data for analysis
	The implementing agency establishes the reliability of the assessment instruments and procedures
Reporting and Analysis	The implementing agency carries out the data analysis
	The draft reports are prepared by the implementing agency and reviewed by the steering committee
	The final reports are prepared by the implementing agency and are disseminated by the appropriate authority
Use of Results	The DoE and other relevant stakeholders review the results and determine the appropriate course of action

Source: Data adapted from Kellaghan and Greaney, 2008

Further, assessment design elements should also be tabulated into a project plan.[12] An excerpt from a sample National Assessment Project Plan is highlighted as Appendix A.

The following sections describe some effective practices that the state government can use to implement each step of their mandate. There are four key aspects to assessment design that a state must consider before a testing framework is created:

1. Class of assessment

Through a state assessment, the government will receive information about the knowledge and skills of students at certain points of their schooling trajectories. The state government must decide which points these should be and whether they should only be for select class or age levels.

In a country like India, where students do not all enter school at the same age, the argument to test at a specific class level is stronger. The state government will have several change-over points, i.e. primary (e.g. Class 3), from primary to upper-primary (e.g. Class 5 or Class 6) and from upper-primary to secondary (e.g. Class 8), that they may want to monitor specifically.

Information collected in early classes (pre-primary, Class 1 and Class 2) can be used to introduce remedial measures. However, it is important to consider that students at these levels require much more personalized testing through oral or one-on-one examination as they may not be able to comprehend written instructions.

Target classes for national assessments have varied from country to country. In the United States, student achievement levels are assessed at Classes 4, 8 and 12; in Colombia, at Classes 3, 5, 7 and 9; in Uruguay, at preschool and at Classes 1, 2 and 6; and in Sri Lanka, at Classes 4, 8 and 10.[13] Some states in India, such as Gujarat and Madhya Pradesh, choose to test each primary and upper-primary class (Classes 1 to 8) each year.

2. Frequency of assessment

A large-scale assessment must be repeated in equal and timely intervals to ensure that the impact of assessment-based interventions can be measured. The gap between consecutive assessments must be significant enough to let such interventions show change. While most international surveys are conducted every four or five years, national or sub-national assessments around the world take place annually or every two years.

The purpose of the assessment will guide its frequency. If the aim of an assessment is to hold teachers, schools and even students accountable for their learning, testing should be carried out every year. Further, because such an assessment focuses on the performance of individuals, as well as performance at the system level, all or most students in the education system should be assessed. Chile and England are examples of such a system.

Contrastingly, if the purpose of an assessment is only to provide information on the performance of the system as a whole, an assessment of a sample of students in a particular curriculum area every three to five years would be adequate. More frequent assessments would be unlikely to register change because education systems do not evolve rapidly. Assessments conducted too frequently would limit the impact of the results, as well as incur unnecessary costs.[14]

3. Population of assessment

A state government must consider whether to assess a sample of the students in a chosen class, or to assess the census of all students in the class. Most international assessments such as the Programme for International Student Assessment (PISA) use the sample-based approach as it reduces the cost of test administration, cuts down the time required to analyze and report results and allows for greater monitoring of field operations. However several regional assessments within countries, such as Minas Gerais, Parana and São Paulo in Brazil, Bogotá in Colombia and Aguascalientes in Mexico use the census method to foster accountability within each school and to provide individualized feedback to each principal, teacher and student.

Several advantages and disadvantages of using the census method to hold schools accountable have been noted in Appendix J.[15]

4. Result compilation

The way in which results will be described should be a consideration at the test development stage. The structure of result compilation will guide the reporting and use of assessment data. Most national assessments have several sets of questions, such that each student only responds to a fraction of the total number of test items. Though administration of such assessments is more complex, it increases the overall test coverage of the curriculum without making individual testing time too long.

Globally, assessments use Item Response Theory (IRT) to report results on a scale that showcases the ability of children vis-à-vis the difficulty of the questions. Cycle three of the NAS has also deployed this technology. An additional benefit of this methodology is that it allows authorities to compare results and check for improvements, across years, as the scale remains constant.

Result compilation also needs to bring forward students' level of subject matter knowledge or the actual skills that students have acquired. Increasingly, national assessments seek to report results in ways that specify what students know and do not know and that identify strengths and weaknesses in their knowledge and skills. This approach involves matching student scores with descriptions of the tasks they are able to do, e.g. "can read at a specified level of comprehension" or "can carry out basic mathematical operations". Performances may be categorized in various ways, e.g. "satisfactory" or "unsatisfactory"; "basic," "proficient," or "advanced" and the proportion of students achieving at each level determined. Matching student scores to performance levels is a complex task involving the judgment of curriculum experts and statistical analysts.[16]

In India, EI has developed a similar 'Scale Anchoring' technique, which clearly shows which concepts or topics are understood only by students performing at higher levels and which topics other students also understand. This technique allows certain topics to 'anchor' at percentile levels of performance, e.g. 25, 50, 75 and 90th percentile and is useful in trying to remediate in a consistent step-by-step manner.

Assessment Framework

Once the key testing design principles for a large-scale assessment have been finalized, an assessment framework is prepared to clarify in detail what is being assessed, how it is being assessed and why it is being assessed.[17] Mullis et al (2006) suggest that a framework may:

- Describe the statement of purpose that guides the rationale for the assessment and specifies what should be measured in terms of knowledge, skills and other attributes.
- Identify and describe various performances or behaviours that will reveal those constructs by specifying number of characteristic tasks or variables to be used in developing the assessment and how those performances are used to assess student performance.[18]

The cognitive assessment framework can be based on:

- Content analysis at a particular class level of what students are expected to have learned as a result of exposure to a prescribed or intended curriculum, or
- The expected level of literacy and numeracy for a certain class level

As an example, the International Association for the Evaluation of Educational Achievement (IEA) has conducted the Trends in International Mathematics and Science Study (TIMSS) for 20 years. The study measures trends in Maths and Science achievement for Classes 4 and 8 students around the globe on a four-year cycle. The TIMSS 2015 Mathematics assessment framework is organised around two dimensions:

- Content dimension, specifying the subject matter to be assessed
- Cognitive dimension, specifying the thinking processes to be assessed[19]

The content domains differ for Classes 4 and 8, reflecting the Maths widely taught in each class (Table 2 and 3). Assessment topics are detailed within each content domain are also detailed (Table 4).

Further information about TIMSS is provided in Appendix B. Three other key international assessments are also detailed in the Appendices. The Progress in International Reading Literacy Study (PIRLS) is also discussed in Appendix B; the Programme for International Student Assessment (PISA) is discussed in Appendix C; the Early Grade Reading Assessment (EGRA) is discussed in Appendix D. Lastly, the Indian NAS is showcased in Appendix E

Table 2:
Target Percentages of the TIMSS 2015 Maths Assessment Devoted to the Content Domain at the Fourth Grade

Content Domain	Fourth Grade (%)
Number	50%
Geometric Shapes and Measures	35%
Data Display	15%

Source: Table reproduced from Mullis, I.V.S. & Martin, 2013

Table 3:
Target Percentages of the TIMSS 2015 Maths Assessment Devoted to the Content Domain at the Eighth Grade

Content Domain	Eighth Grade (%)
Number	30%
Algebra	30%
Geometry	20%
Data and Chance	20%

Source: Table reproduced from Mullis, I.V.S. & Martin, 2013

Table 4:
Target Percentages of the TIMSS 2015 Maths Assessment Devoted to the Cognitive Domains at the Fourth Grade

Cognitive Domain	Fourth Grade (%)	Eighth Grade (%)
Knowing	40%	35%
Applying	40%	40%
Reasoning	20%	25%

Source: Table reproduced from Mullis, I.V.S. & Martin, 2013

Test Items

Test development is a scientific process and requires the collaboration of several partners. The process begins with the structure of the table of specifications, a document that guides test development, analysis and report writing. It describes the data that must be collected, defines the test length and specifies the proportion of items in a test that will address the various aspects of a curriculum. Greaney and Kellaghan (2012) find that a good blueprint should indicate the following:

- The proportion of test items in the final form that address each curriculum area, e.g. Maths, Science and Language.
- The proportion of items within a curriculum area that assess different skills, e.g. in Maths—number, measurement, space and pattern; in writing—ideas, content knowledge, structure, style, vocabulary, spelling and grammar.
- The proportion of items that address different cognitive processing skills such as knowledge or recall, interpretation or reflection.

- The proportion of multiple-choice and open-ended items.
- The proportion of items devoted to stimulus texts of different kinds such as narrative, expository, procedural and argumentative in reading, or tables, charts and diagrams in Maths.[20]

Though an external technical advisor can take the lead in the test development process at the outset, it is imperative that the state agencies also build their knowledge and expertise simultaneously, to be able to take on the process independently at a later stage. As an example, the EI test development programme is showcased (Table 5).

When EI conducted the Student Learning Survey in 2010, they undertook the following steps to construct their items:

- 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, i.e. the difference in the curricula across these states class-wise.
- Systematic study of the National Curriculum Framework, the focus group documents, the minimum levels of learning (MLLs) and existing research on student learning and pedagogy in India.
- National level workshops were carried out with subject experts and assessment experts to finalize the competencies and development of items.
- Changes to the main tests' design 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 and Azim Premji Foundation.
- Three sets of questions were developed for each test and adapted in three 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 harmonization of the versions across the languages.
- Pilots were done in three states and feedback was taken from teachers. The papers were also analyzed for test and item characteristics and fine-tuned.
- The tests were finally carried out in 30 locations assessing 24,600 students.[21]

For a state assessment, it is significantly easier to finalize the competencies to be tested, as they will match the state curriculum and MLLs. At this time, the state can bring together subject matter experts, including teachers, academics and representatives from DIETs and the SCERT to begin item development.

Table 5:
Test Development Process, EI Student Learning Study

The Test Development Stages	
Phase 1: Information Generation Based on Apriori Hypothesis	Defining the Focus Group to be Tested
	Detailed Study of Textbook/Syllabi
	Study of Existing Norms and Standards
	Comparative Study of Other Curricula and Tests
	Item Generation and Development
Phase 2: Item Ordering, Paper Development, Pre-testing for Face Validity and Cognitive Debriefing	Pre-test Instrument Development
	Test Design
	Test Design and Paper Revision
Phase 3: Final Instrument Development and Psychometric Validation	Pilot – Field Tests
	Statistical and Qualitative Analysis
	Fine-tuning Papers, Scorecards etc.
	Translations and Validations
	Linguistic Harmonization – Iterative
	Final Instrument

Source: Table reproduced from EI Student Learning Study, 2010

From EI's experience in question paper design in India, it is known that students, especially in Classes 1 to 5 find it difficult to read questions on their own. Further, the test items need to ascertain students' exact learning levels whist also reaching the expected level of curriculum. To address this, the EI Student Learning Study question paper format was as follows:

- **Group oral:** questions read out aloud by the evaluator, students respond by writing an answer on their paper
- **Written:** questions read and answered by students themselves, including multiple choice, close captioned and free response
- **Individual oral:** for language, answers captured by evaluator[22]

Though both multiple choice and free response items were tested, they were of seven different types:

- Questions check for deeper understanding of concepts
- Questions check for learning that is straightforward or text book-ish
- Questions check for application of concepts
- Questions check for reasoning ability
- Passage questions
- Questions check for holistic language learning and basic writing ability
- International benchmarking questions[23]

Pre-testing or pilot testing of items is an essential element of test development. A pre-test is administered to students with the same characteristics as those who will be taking the final test. Schools of different sizes, in different areas, with students from varying socioeconomic backgrounds should be included. Students who are part of the pre-test sample do not take the final assessment. The role of the pretest is to ascertain whether:

- The test is of the right length
- The test items are of the right difficulty
- Each item has been worded and presented clearly
- Any item in the test is biased towards a section of the population
- The answer format is understood by participants

Lastly, international best practice emphasizes that all test items have model answers detailed at the time of design, especially for free-response items.

Supporting Questionnaires

Most large-scale assessments collect information on student, school and home factors that are considered relevant to student achievement. This would include information on student gender and educational history, including grade repetition; resources in schools, including the availability of textbooks; level of teacher education and qualifications; and socioeconomic status of students' families. International practice is to normally collect this information in questionnaires administered to students, teachers, principals and sometimes to parents at the same time as the assessment instruments are administered.[24]

Identification of contextual factors related to student achievement can help identify manipulable variables, which are, factors that can be altered by policymakers, such as regulations about the time allocated to curriculum areas, textbook provision and class size.[25] Because resources are invariably limited, questionnaires need to be concise and highly relevant.[26]

As an example, in the EI Student Learning Study, background factors related to the student, school principal, teachers and schools were collected (Table 6). Analysis was then conducted to yield insights on relationships, if any, between these variables and student performance. A further detailed questionnaire has been described in Appendix K.

Test developers should also be alerted about some of the challenges of administering such questionnaires and should design the structure accordingly. Anderson and Morgan (2008) describe these challenges as:

- Students may be too young to fill in a questionnaire reliably or accurately.
- Lack of resources may limit the administration of questionnaires to a small group, such as teachers or head teachers, rather than to thousands of students.
- Many parents may be illiterate or unreliable in returning questionnaires.
- Teachers and principals may not be motivated to fill in a long questionnaire, or they may feel too threatened to answer questions honestly.[27]

The School, Teacher and Pupil Questionnaires utilized in the NAS are attached as Appendix F.

Table 6:

Sample Supporting Questionnaire

Characteristic	Variable
School Characteristics	Type of school Multi-grade classroom Infrastructure Material and equipment Hours per year Mid-day meal programme
Head Teacher Background	Gender Age Academic qualification Teacher training qualification Years of experience Type of service
Teacher Background	Gender Age Academic qualification Teacher training qualification Years of experience Type of service Perception on discipline
Student Background	Gender Age Socioeconomic background Parental occupation Tuitions
Student Perceptions	Perception about the school Liking for the subject Use of school library Student reading habits Perception about own academic performance

Source: Table reproduced from EI Student Learning Study, 2010

Section Three

Assessment Administration

ASSESSMENT ADMINISTRATION

Well-planned administration is key to ensuring that the assessment is conducted in a standardized manner. This section describes the main elements of an external assessment, including school preparation, training of administrators, creating an administration manual and recording results.

Self Assessment

Some countries preface their national assessments with a self-administered section that principals and teachers conduct on their own. A few states in India, including Karnataka, Gujarat and Madhya Pradesh, have a similar component in their annual structure.

A self-assessment is usually conducted as a baseline or midline examination, prior to the external assessment. It allows teachers to take stock of student achievement in their classroom on a common set of questions and plan their year, including student remediation, accordingly. In India, a self-assessment can also act as Summative Assessment 1 in the CCE structure.

Test Administrators

The choice of test administrator is extremely important to ensure the validity of testing conditions and reliability of results. Faulty test administration tends to be the most common source of error in a national assessment. Therefore, particular attention should be paid to selecting, training and supervising test and questionnaire administrators.

As per Anderson and Morgan (2008), potential administrators should have the following characteristics:

- Good organisational and communication skills
- Experience working in schools and/or relevant experience in conducting large-scale assessments
- Reliability and ability and willingness to follow instructions precisely[28]

Howie and Acana (2012) suggest that all administrators, regardless of their background, should attend a training session that explains the purpose of the test and their role in its administration. They should understand the importance of following the instruction manual, practice administering the test and clarify questions about the procedures outlined in the manual. They should also be supervised for some part of the day, through a random check.[29]

Anderson and Morgan (2008) have put forword several choices for test administrator:

1. **School inspectors** - The Block and Cluster Education Officers can take on the role of test administration. They can bring in their existing knowledge of school background factors, which will ensure that supporting questionnaires are recorded correctly. However, if the inspectors see test administration as an additional task that is outside their job description, they may not be motivated to do the job properly.

   School inspectors require significant administration training to ensure that they follow instructions reliably. Additionally, senior officials need to urge them to consider this an important task that will have a resultant outcome on the training, remediation structures and resource allocation in their schools

2. **External administrators -** External administrators are used in some national assessments. Anderson and Morgan (2008) suggest that they are people who can follow instructions precisely, have the time and resources to do the task properly and have no particular interest in the outcome of the test other than to administer it correctly.

   External administrators were appointed for the EI Student Learning Study (SLS) in 2010. The structure for this recruitment was as follows:

a. Participating states were divided into five zones, each managed by a Zonal Manager. The SLS was coordinated by a Project Manager

b. 21 State Coordinators and 60 District Coordinators were recruited, one for each state and district targeted by the study.

c. These coordinators, together, carried out recruitment of evaluators. 20 to 30 evaluators were required for one to two weeks, in each participating district.

d. Evaluators were recruited from colleges of education, social work and other social sciences. 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 test was administered and evaluators were selected based on their:

- Performance in the test .
- Voice assessment: They were asked 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.

e. Two to four master trainers were appointed, responsible for the training of evaluators and ensuring quality of test administration in each state.

f. Two 3-day workshops were conducted to train the master trainers.

g. 2-day workshops were held for each evaluator in conducting oral and written tests, assimilating scorecards, invigilation etc.

h. Training manuals were created and documented for each role.

i. The selected evaluators were given a stipend and a certificate for participation at the end of the study.[30]

3. **Teachers** - Teachers can administer the assessment directly to students in their class, which can create an inviting 'low-stakes' environment and boost student morale. The key concern, however, is that teachers may, deliberately or unintentionally, offer assistance to the students. Additionally, if the teachers are also test scorers, there is not enough division of power to ensure validity of results.[31]

Test Scorers

Test scorers support test administrators in checking student responses. If possible, it is recommended that large-scale assessments use Optical Mark Recognition (OMR) sheets to record and score multiple-choice questions and only utilize test scorers for free responses. Alternatively, the AP RESt study in India has also used 'transcription' sheets, which are low-cost and utilize double entry software to accurately record student responses. In either case, test scorers require adequate knowledge of subject matter being tested. In most countries, teachers, members of DoE staff or university students take on this role.

School Preparation for Testing

The state assessment should be mandatory for all participating government schools, to ensure validity of data. Schools should be informed of their participation as soon as they are selected and in case of a census assessment, at the start of the academic year. This information should be imparted in the form of a letter or through hosting a seminar, with

the purpose of the assessment, tentative dates and administration procedures highlighted.

EI began the school preparation process for the SLS with a school verification visit, to check the physical location, travel modes, school timings, enrolment, classroom size and structure. This ensured that the school details provided by the district, or outlined in the District Information System for Education (DISE) data, were accurate. It also allowed EI to prepare for special administration instructions in schools where there was a paucity of space or resources.

In Karnataka, as the Karnataka State Quality Assessment and Accreditation Council (KSQAAC) pilot-tested their student assessment through the school preparation process, they realized that students were unable to take multiple lengthy examinations in one day. Their fatigue was impacting results in the last assessments of the day. Through this process and by collecting school feedback, KSQAAC lengthened the structure of their assessment to three days.

All international best practice recommends that principals and teachers in participating schools should know the purpose of the assessment. They should be told that their particular schools and classes have been selected to help gain information about what students do and do not know. They should know that individual schools or classes are not being judged. Principals and teachers in participating schools should also be told that all test data and questionnaire responses will be treated as confidential.[32]

Test Administration Manual

A standardised manual must guide test administration so that all students take the test under the same conditions. Anderson and Morgan (2008) state that the manual should specify the exact conditions under which a test must be conducted, including preparation requirements and procedures for ensuring test security. They should emphasize that students taking the test must work through the same practice questions and receive the same instructions about how to show their answers. All must be given the same amount of time to do the test with the same degree of supervision.[33]

The manual should be read and reviewed by the test administrator, principal and all teachers and staff members involved in the assessment process. A workshop should be conducted for all principals in a block, advising them of the assessment procedures and ensuring that they convey the same details to their teachers. If teachers are a part of test administration procedures, they should attend separate training sessions, which outline their role and also allow them to practice their duties with master trainers.

Further, the manual should be supplemented with a checklist that can be used to track student answer booklets and maximise security. Some of the other questions an administration manual should answer have been detailed in Exhibit 1.

Exhibit 1:

Questions an Administration Manual should answer

1. What is the test for?

- Brief explanation of the purpose of the test and the way the data will be used

2. Which tests are given, which students are tested and when are they tested?

- Which tests are being administered in the school
- Which students should take each test
- Dates and times of test administration
- Order of administration of tests
- Length of time of administration of each test
- Any required breaks between test administrations
- Options for flexibility in the administration schedule

3. What test materials are needed?

- List of all the test materials that are supplied
- Quantities of each test material supplied, such as one per student or one per teacher
- List of any materials the school needs to provide, such as pencils and erasers

4. How should the room be set up for the test?

- Physical facilities the school needs to provide, such as desks and chairs
- Resources that might assist students should be removed from the room or covered up, such as charts of multiplication tables or posters displaying grammatical rules

5. What preparation is required?

- How the principal or head teacher might motivate staff members and students to support the administration of the test before the actual administration
- What information the test administrator might require, such as a list of names
- How test booklets might need to be sorted, numbered, or named
- How student groups might need to be organised for testing

6. How should the test be conducted?

- How students should write their name on booklets and record background information on the front cover
- When and how the administrator should check that students have correctly recorded the information on the front cover of the test booklet
- How the practice questions should be administered and explained
- What instructions the students should receive about the test
- What level of support the administrator can offer during the test
- How long students have to complete the test
- What conditions the administrator needs to maintain during the test
- Who should be allowed into the room during test administration

7. How should test materials be stored?

- Procedures to ensure the security of the test materials before, during and after the test

8. Who can be contacted for help?

- Contact details for people who can assist with problems or provide additional info.

Source: Exhibit reproduced from Anderson and Morgan, 2008

APF has also highlighted several cultural practices about test administration that should be kept in mind. These are listed below:

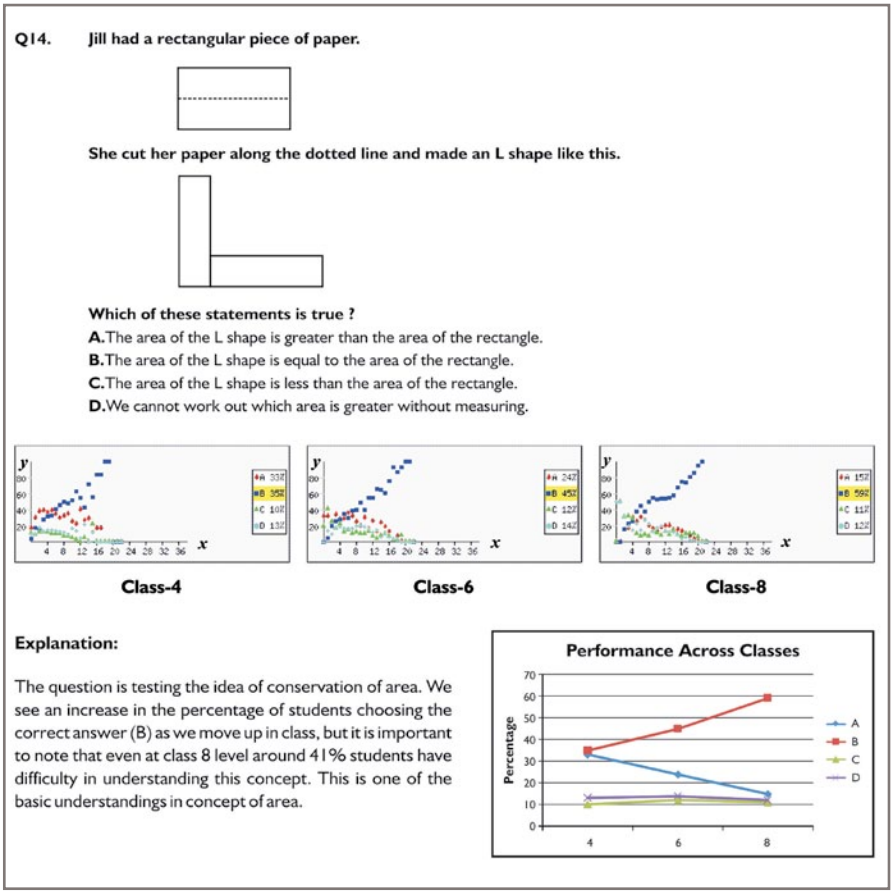
- Every child being assessed should be treated with dignity: calling children by random names, physical proximity, coddling, etc. should be discouraged.
- Children should feel comfortable: conducting a few fun activities with children to build rapport and familiarize them with the assessors can help in breaking the ice with children.
- All required details of schools and learners should be filled and collected accurately.
- For written papers instructions should be read aloud and clearly to the children.
- For oral papers, the assessors must ensure that children do not feel threatened in their presence and are made comfortable.[34]

Test Scoring and Data Recording

Large-scale assessments typically employ the use of three types of test items – multiple-choice, closed-constructed questions and free response. The fourth type of test item, the essay, is cumbersome and expensive to score reliably in a census assessment.

Items that require hand-scoring cost more and take more time and can delay the publication of a report. The more complex the scoring guides, the greater the costs. Essays

Exhibit 2:
Example Student Misconception Analysis



Source: Exhibit reproduced from EI, Quality Education Study, 2006

and extended-response items tend to cost most. Multiple-choice items cost less to score but are more expensive than other item types to construct. Multiple-choice items are usually scored as correct or incorrect by the data analysis software.

Scanning is most economical for large-scale testing. It requires special equipment and sometimes technical backup support. If data are being scanned, one must ensure that all responses are recorded.[35] Valuable diagnostic information about student performance can be obtained by recording each option. EI, through the 2007 Quality Education Study, has published the Student Misconception and Common Error Report (Exhibit 2), which features topics with weak or incomplete conceptual understanding through analysis of wrong answer choices.

Greaney and Kellaghan (2012) suggest that in planning data recording, the state should calculate the amount of time necessary to enter and verify data for each test, such as one Maths test booklet and one Language test booklet and each questionnaire, such as student and teacher questionnaires. This should provide with an estimate of the amount of time that will be needed to enter or type and verify all the data. This estimate will give a rough guide to how many data entry personnel will be needed to complete the task on time.

After determining how many staff members will be needed, one computer should be provided for each data entry person, as well as one for the supervisor. Ideally, computers should be linked to a network. The authors also comment that some national assessment teams use custom software (such as the International Association for the Evaluation of Educational Achievement's WinDem or EpiData) for data entry; others use database packages such as Access and Excel.[36]

Section Four

Assessment Reporting and Use of Results

ASSESSMENT REPORTING AND USE OF RESULTS

It is vital that assessment results are analysed and reported in a timely manner to all stakeholders. This reporting and dissemination can take various forms, including flash-statistics, summary of analysis, reports at various units of collation and conferences. This section also details the use of assessment results in policy making, resource planning, teacher training and raising public awareness.

Result Analysis

Data analysis requires a team with proven competencies in statistics, including psychometrics according to Kellaghan and Greaney (2009). Additionally, the authors state that it can take a considerable amount of time to select appropriate hardware and specialized software; get release of government, donor or other funds; order equipment and software; and have it installed and operational. Exhibit 3 showcases steps in the data

Exhibit 3:

Steps in the Data Analysis Process

1. Secure competent statistical services.
2. Prepare a codebook with specific directions for preparing data for analysis.
3. Check and clean data to remove errors (e.g. relating to numbers, out-of-range scores and mismatches between data collected at different levels).
4. Calculate sampling errors, taking into account complexities in the sample, such as stratification and clustering.
5. Weight data so that the contribution of the various sectors of the sample to aggregate achievement scores reflects their proportions in the target population.
6. Identify the percentage of students who met defined acceptable levels or standards.
7. Analyze assessment data to identify factors that might account for variation in student achievement levels to help inform policy making.
8. Analyze results by curriculum domain. Provide information on the subdomains of a curriculum area (e.g. aspects of Maths, reading).
9. Recognize that a variety of measurement, curricular and social factors may account for student performance.

Source: Data adapted from Kellaghan and Greaney, 2008

analysis process. Many national assessments, including the NAS, have released tenders to engage an external party to conduct this process. Appendix G captures a snapshot of the NAS tender.

Specific recommendations for the data analysis process have been detailed in Exhibit 3. Kellaghan and Greaney (2010) have also identified common errors that assessment analysis teams may commit:

- Using inappropriate statistical analyses, including failing to weight sample data in the analysis.
- Basing results on small numbers, e.g. a small sample of teachers might have responded to a particular question.
- Contrasting student performance in different curriculum areas and claiming that students are doing better in one area on the basis of mean score differences.
- Failing to emphasize the arbitrary nature of selected test score cutoff points, such as mastery versus non-mastery, pass versus fail, dichotomizing results and failing to recognize the wide range of test scores in a group.
- Not reporting standard errors associated with individual statistics.
- Computing and publicizing school rankings on the basis of achievement test results without taking into account key contextual factors that contribute to the ranking. Different rankings may emerge when school performances are compared using unadjusted performance scores, scores adjusted for contextual factors and scores adjusted for earlier achievement.

- Inferring causation where it might not be justified (e.g. attributing differences in learning achievement to one variable, such as private school administration or class size).
- Comparing test results over two time periods even though non-equivalent test items were used.
- Comparing test results over two time periods without reporting the extent to which important background conditions, such as curriculum, enrolment, household income or level of civil strife might have changed in the interim. Although most education-related variables tend not to change rapidly over three to four years, some countries have introduced policies that have resulted in major changes in enrolment. For instance, following the abolition of school fees in Malawi and Uganda, the number of students enrolling in schools greatly increased.
- Limiting analysis to a listing of mean scores of geographical or administrative regions.[37]

Report Writing and Dissemination of Findings

International best practice suggests that reporting on an assessment must include details on findings and testing procedures, to provide context to readers. Kellaghan, Greaney and Murray (2009) suggest that the main report of a state assessment contain the following sections:

1. **Context of the assessment –** This section relates to the importance of capturing the level of student learning and the vision with which this assessment has been sanctioned.
2. **Objectives of the assessment –** The objectives highlight the purpose of the evidence, i.e. the specific evidence the assessment aims to gain and what outcomes it hopes to achieve.
3. **Framework for the assessment –** The framework highlights the knowledge and skills that were tested and the rubric on which the students were marked. This framework also describes the instruments and item types used in the assessment and provides background for analysis of answers.
4. **Procedures in administration –** The test administration process, including selection of schools, background of assessors, in-school procedures and data input.
5. **Description of achievement –** Increasingly, proficiency levels using scale anchoring are used to present the results of assessments as discussed earlier. The levels may be labeled, e.g. in quartiles and the proportion of students achieving at each level identified.

   The findings of a state assessment should be presented so they are relevant to policymakers' and decision-makers' needs in addressing policy problems constructively. Although policymakers may generally prefer summary statistics, reporting only a single index of achievement will most likely miss important information and limit the basis for action following the assessment.

6. **Correlates of assessment –** The supporting questionnaires that accompany a state assessment provide a significant amount of data about background factors that may impact the achievement of subgroups within the population.

7. **Changes in assessment over time –** The purpose of repeating the same, reliable assessment over time is that it can showcase trends in student learning. These trends can be followed not just at a summary level, but also for specific strands of the curriculum, geographical areas or the like. In essence, if any specific action has been taken to improve learning in between assessments, its impact can be tracked directly.[38]

Though this type of report is the key post-assessment publication, several other dissemination structures should be employed to reach out to a broader audience beyond policymakers. Pérez (2006) has identified the following general principles regarding the communication of research findings that are applicable to national assessment findings:

- Use simple language, preferably in attractive media products such as videos.
- Clearly identify stakeholders and tailor events and products to their needs.
- Recruit public and credible leaders as advocates.
- Disseminate information to mass media. Events should be well advertised.
- Use slogans and simple messages that are readily understood. For example, a statement such as "an eight-year-old child should be able to read a 60-word story in one minute and answer three questions about its content", illustrates what a standard means.
- Back up all broadcast or large-audience dissemination materials (including PowerPoint presentations) with supporting technical information.[39]

Several other internationally utilized dissemination structures, detailed in Kellaghan, Greaney and Murray (2009) include:

1. **Briefings for ministers and policy personnel –** Ministers and their senior officials often do not have time to read full reports, but they do need to be aware of key findings and to issues that the media, parliament, or stakeholders in the education system may raise when the report of an assessment is published. Because ministers tend to get numerous documents to read on a daily basis, a briefing note must be short and to the point. Particular attention needs to be paid to how differential outcomes for subpopulations are reported and interpreted.

2. **Publishing summary reports –** Classroom teachers form the primary readership of non-technical summary reports. These should be very brief, and contain information on student performance, examples of test items and trends in learning. Other interest groups that authors suggest to be usually interested in specific summary results include teacher unions, community leaders, employers, businesses and donor agencies.

3. **Publishing technical reports –** Technical reports are a crucial element of a national assessment because they provide members of the research and scientific communities with detailed information about the assessment that allows them to evaluate it critically. Technical reports also act as a record of the activities involved

in the assessment, which is needed to implement future cycles of an assessment.

4. **Publishing thematic reports –** Thematic reports explore aspects of the findings of an assessment related to a specific theme that are not addressed in detail in the main report. A thematic report could analyze error patterns in students' responses to particular aspects of the curriculum or to sets of items in an achievement test e.g. a student misconception report, as per EI SLS. Such analyses can help identify where a curriculum needs to be reformed or instruction needs to be strengthened.

5. **Securing media reports –** The media can provide an inexpensive way of disseminating the main messages of a national assessment to the public at large. For example, the results of an international assessment in South America carried out by the Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación were publicized through a video shown on television throughout the continent. Research in Peru shows that videos were much more effective than lectures or PowerPoint presentations in dialogue with stakeholders on educational policy.[40]

6. **Issuing press releases –** The press release must be drafted keeping the audience in mind, to ensure the right level of technical information. Preparing a press release helps reduce, but does not eliminate, the tendency of reporters to oversimplify assessment findings. Appendix H shows the national press release issued by National Assessment of Educational Progress (NAEP) in 2013 for the Trial Urban District Assessment (TUDA) in Maths and reading.

7. **Holding press conferences –** Press conferences can reach out to the media and general public. However, it must be well programmed, with speakers coming together to deliver a common message. The popular media in many countries have a tendency to provide simplistic explanations of complex issues, such as causes for falling or rising standards of student achievement. It is important that the press conference be used to correct unwarranted conclusions, such as laying the blame for poor results on a single cause or group.

8. **Posting website reports –** Several countries have published a form of their results online. Chile publishes detailed results on its national assessment web site.[41] The Irish Department of Education and Science presents summary findings of national assessments on its official web site.[42] In India, Madhya Pradesh publishes student-level results of the Pratibha Parv on its state education portal.[43]

9. **Making assessment data available –** Actual assessment data are an often-neglected asset; however, a variety of agencies or individuals might have an interest in carrying out secondary analyses of the data. Barriers to use by individuals or agencies other than the national assessment team, while largely technical, include issues of privacy and confidentiality.

10. **Targeted workshops –** Conferences, workshops and seminars provide mechanisms to advertise the availability of assessment results to key stakeholders in a variety of ways. They provide the opportunity to reach consensus on the meaning of key findings and on the steps that need to be taken to remedy any identified problems.

Use of Results

The information that an assessment provides about student learning should inform the decision-making process in the state. This can be done in three ways, as described by Kellaghan, Greaney and Murray (2009):

1. **Policy and education management -** The state assessment provides information about student achievement, school and teacher resources and trends across the population. Policymakers can use this to:

   - Formulate general policy and assist in decision-making
   - Plan for financial resources
   - Track implementation success of existing plans
   - Set standards
   - Provide additional resources to schools
   - Support curriculum revision
   - Revise textbooks

   Appendix I outlines a list of the use of assessment results by several countries.

2. **Teaching –** Teacher educators who focus on enhancing teachers' professional development can utilize achievement results. Karnataka, detailed in a case study later in this document, specifies roles for state, district, block and cluster officers and also for principals.

3. **Public awareness –** Assessment findings may fail in their purpose to inform the public because the reports are too technical. Here, the state can utilize the structures of SMCs, constituted in all government and government-aided schools. SMCs are mandated to monitor the learning environment of the school and should be briefed about school and student progress, so that they can provide this information to the parent community. The SMC can also organise parent-teacher meetings to discuss individual student progress.[44]

Section Five

Budget and Cost Structures

BUDGET AND COST STRUCTURES

This section discusses the cost structures for the periodic implementation of a large-scale assessment. Sample cost-per-child financial numbers from two sources are provided.

Ilon (1996) describe a number of costs involved in the structure of a state assessment. These are outlined in Exhibit 4:

Exhibit 4:

Cost structure of a state assessment

1. **Implementing agency –** Costs will vary depending on whether the agency has the necessary facilities and expertise or needs to upgrade or employ full-time or part-time consultants. The cost of providing facilities and equipment, including computers and software, also needs to be taken into account.

2. **Instrument content and construction –** Options for the selection of the content and

form of assessment should be considered in terms of cost, as well as other factors, such as validity and ease of administration. Multiple-choice items are more expensive to construct than open-ended items but are usually less expensive to score. The cost of translating tests, questionnaires and manuals and of training item writers also needs to be considered.

3. **Numbers of participating schools and students –** A census-based assessment will be more expensive than a sample-based one. Costs increase if reliable data are required for sectors of the system, such as states or provinces. Targeting an age level is likely to be more expensive than targeting a class level because students of any particular age may be spread over a number of classes, requiring additional assessment material and testing sessions.

4. **Administration –** Data collection tends to be the most expensive component of a national assessment. It involves obtaining information from schools in advance of the assessment; designing, printing, packaging and dispatching test materials and questionnaires; and establishing a system to administer instruments. Factors that contribute to overall cost include:

    - The number of schools and students that participate
    - Travel
    - Difficulty in gaining access to schools
    - Accommodation for enumerators
    - The collection and return of completed tests and questionnaires

5. **Scoring, data management and data entry –** Costs will vary according to the number of participating schools, students, teachers and parents; the number of open-ended items; whether items are hand or machine scored; the number of inter-rater reliability studies; and the quality of test administration and scoring.

6. **Analysis –** Analytic costs will depend on the type of assessment procedures used and the availability of technology for scoring and analysis. Although machine scoring is normally considered to be cheaper than hand scoring, this may not be the case in a country or state where cost of labour is low.

7. **Reporting –** Budgeting should take account of the fact that different versions of a report will be required for policymakers, teachers and the general public and of the nature and extent of the report dissemination strategy.

8. **Follow-up activities –** Budgetary provision may have to be made for activities such as in-service teacher training, briefings for curriculum bodies and secondary analyses of the data. Provision may also have to be made to address skill shortages in key professional areas.[45]

Source: Data adapted from Ilon (1996)

Additionally, Kellaghan and Greaney (2008) have compiled the following funding checklist for national or state assessments. The authors suggest that the source of funding for each item, including from within or outside the state budget should be highlighted at the outset:

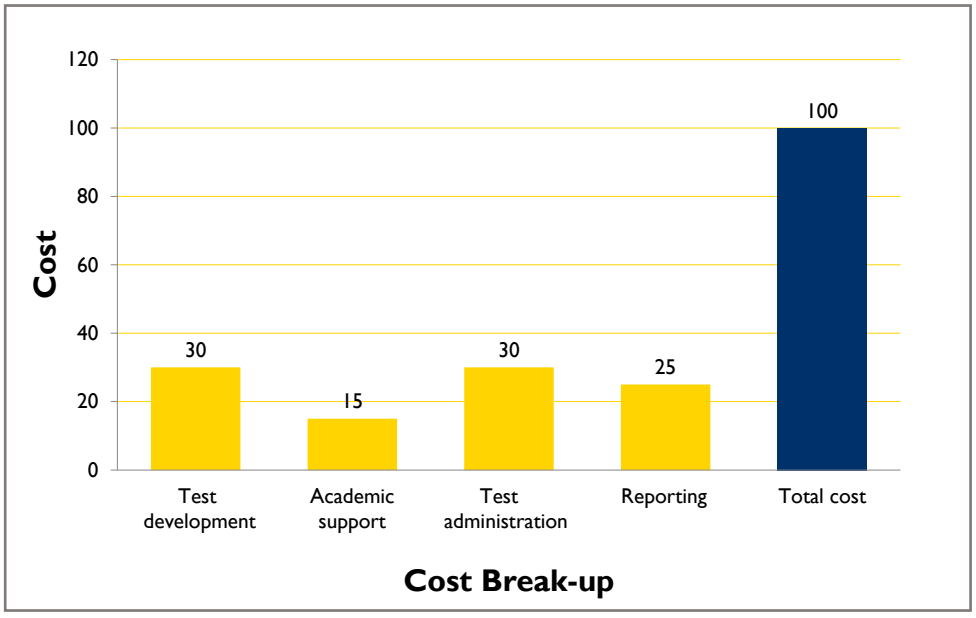
1. Personnel
2. Facilities and equipment

3. Design of assessment framework
4. Instrument design and development
5. Training (e.g. item writing and data gathering)
6. Pilot-testing
7. Translation
8. Printing
9. State level committee
10. Local travel (to schools)
11. Data collection
12. Data scoring (open-ended)
13. Data recording
14. Data processing and cleaning
15. Data analysis
16. Report writing
17. Printing of reports
18. Press release and publicity
19. Conference on results
20. Consumables
21. Communications
22. Follow-up activities

The cost of conducting a large-scale assessment in India is around ₹100 per student, as per estimates by Kaizen PE in 2013. The typical break up of cost is shown in Exhibit 5.

Assuming the cost per student of ₹100, the total cost of conducting assessment statewide will be in the range of 0.3% to 2% of the total education budget of the state (Table 7). Additionally, the budget for the Andhra Pradesh Randomized Evaluation Studies (AP RESt) Assessment is showcased (Table 8).

Exhibit 5:
Break-up of Per-Child Cost of Large-Scale Assessments



Source: Kaizen PE estimates

Table 7:
Estimated Cost of Census State Assessment as a Percentage of Per Child Spend on Education

States	Per Child Spend on Education (in ₹)	Estimated Cost of Assessment (%)
Kerala	37,667	0.3%
Jharkhand	5,669	1.8%
West Bengal	6,954	1.4%
Maharashtra	21,226	0.5%
Himachal Pradesh	29,785	0.3%

Source: PAISA (2012), 2012-13 Budget Estimates, Kaizen PE Calculations

Table 8:
AP RESt Budget

Particular	Per Child Cost in 2014 (in ₹)	Total Children	Total Cost (in ₹)
Printing of Assessment Papers	25	10,000	2,50,000
Stationery	10	10,000	1,00,000
SPOCs Orientation	10	10,000	1,00,000
Evaluators' Orientation	40	10,000	4,00,000
Evaluators' Honoraria	150	10,000	15,00,000
Data Entry	10	10,000	1,00,000
Transportation	15	10,000	1,50,000
Miscellaneous Expenses	10	10,000	1,00,000
Total	**270**	**10,000**	**27,00,000**

Source: Data adapted from Azim Premji Foundation

Table 9:
AP RESt Budget

Particular	Cost Per Child @ 75,000 Children in 2010 (in ₹)	Cost Per Child @ 5,000 Children in 2013 (in ₹)
Printing of Assessment Papers	14	25
Stationery	4	163
SPOCs Orientation	2	25
Evaluators' Orientation	4	25
Evaluators' Honoraria	50	165
Data Entry	7	15
Transportation	25	25
Miscellaneous Expenses	15	25
Total	**121**	**468**

Source: Data adapted from Azim Premji Foundation

Section Six

Case Studies of Large Scale Assessments

CASE STUDY 1: KARNATAKA

The Mandate

The Government of Karnataka constituted the Karnataka School Quality Assessment and Accreditation Council (KSQAAC) as an autonomous body in November 2011.

KSQAAC functions under the guidance of its Governing Council, which is headed by the Minister for Primary and Secondary Education.

History

The Azim Premji Foundation (APF) conducted an assessment-based intervention, the Learning Guarantee Program (LGP) in Karnataka from 2002 to 2006. The LGP was designed on the premise that better assessment would help gauge the real 'level of learning' in government schools and thus develop a stronger accountability structure. APF designed the LGP as a competency-based assessment and implemented it in 1,800 government schools over three years. Schools opted to participate in this assessment and got incentives for good performance.

The Government of Karnataka set-up the Karnataka School Quality Assessment Organisation (KSQAO) to sustain the LGP assessment mandatorily in all schools from 2006 onwards. However, this assessment structure was terminated after one year as the process became high stakes for students and school managements.

Purpose

In 2011, KSQAAC was set up to:

- Assess and provide accreditation to primary and secondary schools across the state (government, private aided and private unaided), considering the school as a wholesome unit.
- Stimulate academic activities in the school, focusing on promotion of quality in the school and in the schooling processes.
- Promote collaboration amongst all stakeholders in the planning and implementation of quality.
- Encourage school staff and School Development and Monitoring Committees (SDMCs) to ensure good quality facilities in school.
- Assess the learning achievement of all students in the school and provide guidance for further improvement.
- Help utilize all academic and physical infrastructure for the improvement of quality.
- Promote and encourage necessary changes, innovation and reform in all aspects of the institution through both self and external assessment.
- Progressive community involvement and accountability amongst staff, management and SDMCs for the betterment of schools.

Assessment at a Glance

Table 10:
KSQAAC Snapshot

Parameter	Description
Structure	A sample of Kannada medium government schools
Scale	1,020 schools across the state, with three upper primary schools and three high schools in every block
Frequency	Annual – different sample of students assessed every year
Classes	3, 5, 7 and 9
Achievements	Kannada, Maths, Science, Social Science, English (Classes 7 and 9), Non-Scholastic
	Competencies are oral and written at the primary level and only written in higher classes
Indicators	Institutional vision and mission
	School physical environment and infrastructure
	Classroom environment and process
	Teaching-learning process and learning achievement
	Teachers professional development
	Community participation

Instruments

The evaluation instrument is competency based and designed by the Karnataka State Council for Educational Research and Training (KSCERT) resource team for overall school

improvement. The instrument is data driven and builds on self and peer assessment. The data from the assessment provides impetus for change by identifying opportunities for staff's professional growth, strenghtening of school and community relations.

Learning standards in each subject from Classes 1 to 10 are identified by subject experts and practicing classroom teachers. Comprehensive learning competencies are then identified based on the learning standards and student achievements in previous classes. Finally, weightage is given to the different components. There are 175 indicators in total, across the following categories:

- Physical Infrastructure: 20%
- Learning Environment (teaching, learning and evaluation processes): 60%
- Leadership: 10%
- Community Participation: 5%
- Innovation: 5%

Test Administration

Selection of schools
Block Education Officers in each of the 284 blocks in the state of Karnataka are asked to select schools for the pilot study, with some guiding criteria.

School self assessment
The assessment and accreditation process begins with school self-assessment. Each headmaster and one senior member from each institution is trained in a two day residential training to assess their school and identify areas of improvement. They are provided with guidebooks, processes, methodologies and the indicators upon which the school will be assessed. The self-assessment format focuses on the details of the institution, including staff pattern, teacher training, physical infrastructure, learning strategies, teaching techniques and methodologies, evaluation structures, use of technology, leadership and community participation.

This process allows the school management to introspect and instigate school improvement. It ensures buy-in and goodwill at the school level for the quality assurance process that the external administrators bring in at a later stage. Lastly, as this is a pilot programme, the self-assessment allows KSQAAC to understand preparations required for the actual assessment.

Peer team assessment
A team of three external peer assessors – one each for Kannada/History, Biology/English and Physical Science/Maths – is sent to each school for a period of three days. In this time, the peer assessors are given a pre-prepared observation template and marking guidelines.

A state-level committee selects the assessors. The minimum requirement for becoming an assessor is to have a university degree in humanities or science with a degree in education and with teaching experience of at least one year. These positions are advertised across print and online media and nearly 8,000 people applied. Most applicants are retired teachers, retired officers and unemployed trained graduates. The structure of their application review is as follows:

- Of the 8,000 applicants, 5,555 appeared for the first written examination. This examination tested two key components – first, the knowledge of the current education system in Karnataka and second, content and academic knowledge.
- Interviews were conducted for the 3,807 candidates who qualified based on the written examination. These interviews were conducted at the divisional level by a committee of DIET nodal officers and two subject matter experts.
- 612 assessors were selected through the interview.

Each assessor receives five days of intensive training on the invigilation structure. This training is delivered by Master Resource Persons at the district level.

Student learning assessment

While the accreditation process viewed the five holistic pillars of school quality with importance, it is key to understand how student learning, specifically, was assessed. All the school and classroom practices and community and management leadership culminate in the improvement of this outcome and hence, this was given a higher weightage. The structure of this assessment was as follows:

- Each subject team, from Classes 1 to 10, is headed by an eminent professor.
- This team identifies learning standards for each subject.
- They then selects competencies to be tested for each class, giving due weightage to the current and prior grade. Previously learnt content is given a weightage of 60% and current content 40%. Key competencies per this weightage are decided in a workshop conducted with practicing school teachers and forms the basis for preparation of question papers.
- Strong subject teachers are invited for a four-day workshop to design question papers on the basis of the competency blueprint. These teachers are oriented to the principles of constructing test items and their importance in assessing student achievement. 30 competencies are identified in each subject and questions are prepared for each. The first 20 questions are multiple-choice items and 10 are framed to test written and comprehension abilities.
- Two teams of subject experts refine and scrutinize every question and prepare a total of five sets of question papers for each assigned subject.
- Optical Mark Reader (OMR) sheets are used to record student answers. Where comprehension or oral answers are required, external assessors issue instructions and fill OMR sheets manually at the end of the assessment.

Reports and Analysis

For this cycle of the KSQAAC, a committee was formed at the district-level DIET, under the chairmanship of the DIET principal and the DDPI. This committee had the nodal officer, one DIET lecturer and one educationist as members. It reviewed all the self-assessment reports submitted by schools and assigned marks as per the criterion fixed by the KSQAAC. The consolidated marks of all schools of the district were sent to KSQAAC, who then compared them with marks awarded by the external assessors.

Accreditation was awarded on the basis of holistic achievement across the five assessed areas. The percentage assigned was as follows:

Table 11:
KSQAAC Rating Structure

Marks	Grade
90.1% and above	A+
80.1% to 90%	A
70.1% to 80%	B+
60.1% to 70%	B
50.1% to 60%	C+
40.1% to 50%	C
Below 40%	D

For student learning results, OMR sheets were scanned to collect raw data, which gave details about the questions attempted by each student. These results were aggregated by subject, school, block, district and state and on other parameters like gender. Inferences were drawn accordingly.

Use of Results

The KSQAAC accreditation report for each school indicating its performance will be made available to all stakeholders. Each school is encouraged to follow a number of steps:

- Identify the competencies of higher and lower achievement in each subject and locate the causes for the same
- Prepare an action plan to enhance the learning of students in each competency
- Conduct action research to initiate remedial measures
- Consider achievement of other schools to find out what encouraged success in particular areas
- Organise staff meetings to further discuss student competency
- All stakeholders, including the SDMC, should analyze how physical infrastructure, teaching-learning strategies, innovation and leadership have influenced student achievement

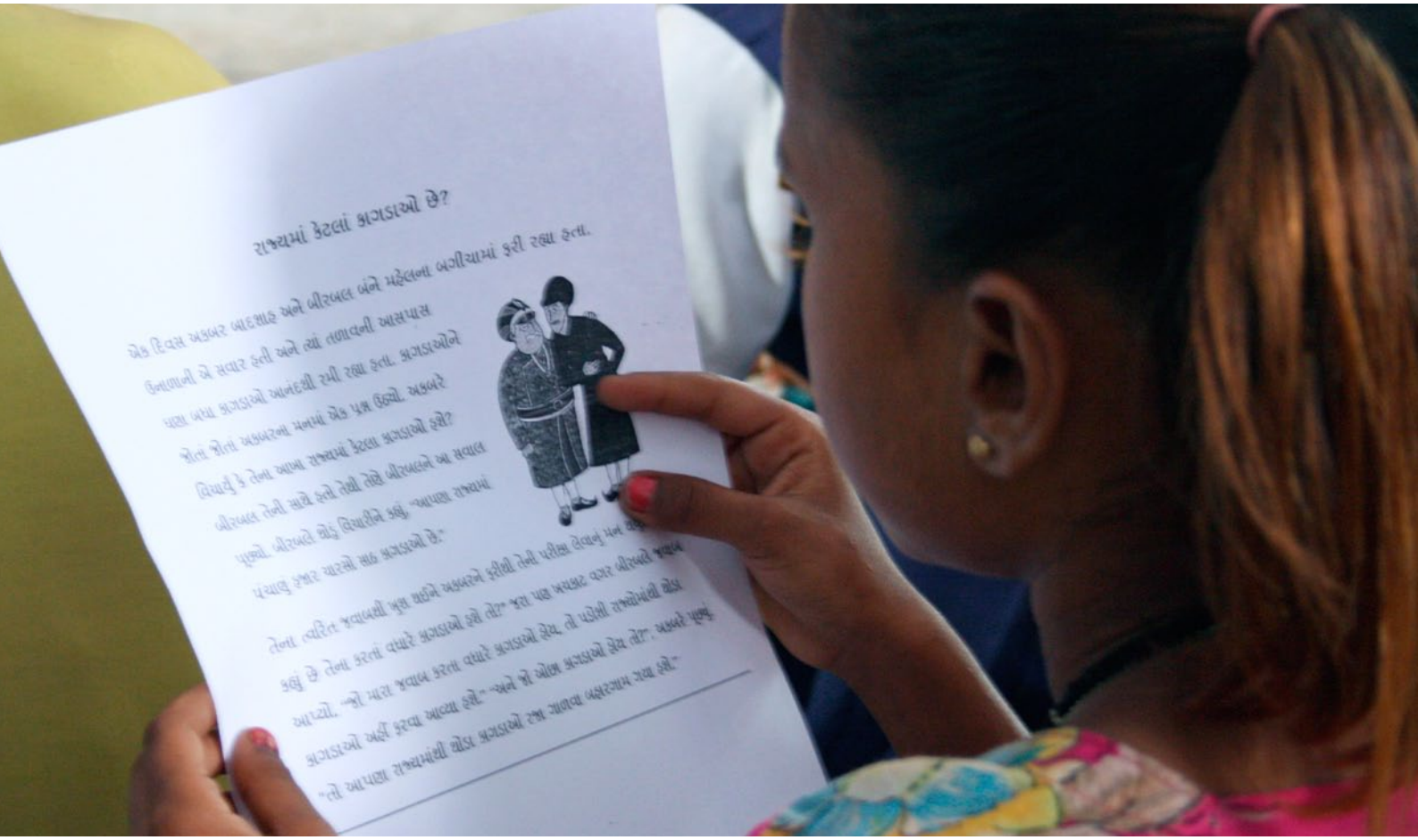
At the block level, field functionaries including the BEO, BRC and CRC are also sent reports, which discuss achievement across the block by school, gender, social group, category, competency, subject and enrolment. The team at the block level is encouraged to:

- Compare their results with those of other blocks in their district and the state
- Identify competencies with lower achievement and understand causes for these
- Compare the performance of schools in different geographical areas to understand how other background factors may have influenced student performance

- Train Block Resource Persons (BRPs) and Cluster Resource Persons (CRPs) to conduct similar analysis at the cluster level and develop and action plan for improvement
- Encourage schools to adopt innovative practices showcased during this process and socialize student performance data
- Every action initiated by the BEO is followed up by the CRPs and made a regular agenda item at cluster and block level meetings

At the district level, officers including the DDPI, DIET nodal officer and Deputy Project Coordinators of the SSA and RMSA receive results. From this, district officers and field functionaries should prepare district level plans, consisting of measures to improve learning competencies and they should guide block level functionaries in implementing these. Furthermore, they should conduct surveys to understand the trainings needs of teachers and include these in the DIET annual plans. Lastly, the district should organise sessions where 'A' grade accredited schools share their experience and provide resource support to others follow similar best practices.

At the state level, Karnataka has set-up Internal Quality Improvement Cells (IQIC) to inspire everyone from the school to the state level. Also the Department of State Educational Research and Training (DSERT) has been mandated to arrange workshops and trainings structures to analyze and identify the most difficult competencies and suggest remedial measures for these.



CASE STUDY 2: GUJARAT

History

Till 2009, the Gujarat government's focus in education was on enrolment. However, there was continued discourse about the quality of primary education. To address this, the Gujarat government launched a programme called Gunotsav ('Celebrating Quality').

In November 2009, the state education department started this programme to evaluate primary education and grade teachers. Gunotsav brings together several state departments other than education, through the assessment process, to create awareness amongst government stakeholders about learning outcomes in the state. The programme feeds into the larger goal of ensuring that Gujarat is ranked among the top three states of the country in terms of student learning outcomes over the next five years.

Purpose

The stated objectives of Gunotsav are as follows:

- Bring awareness in teachers and in the general public for quality education
- Provide educational evaluation and grading of schools and teachers
- Assess quality based on classroom educational work or teaching process followed
- Facilitate improvement in education's quality in primary schools

Assessment at a Glance

Table 12:
Gujarat Gunotsav Snapshot

Parameter	Description
Structure	Classes 2 to 8 in all schools in Gujarat
Scale	The self-evaluation is conducted by 33,922 primary schools, including 32,774 government primary schools, 701 Ashram schools and 447 granted schools
Frequency	The self-evaluation component of the assessment is conducted on an annual basis. Senior government officials also assess 25% of schools in each block each year. Lastly, a sample diagnostic assessment of students in Classes 3, 5, 7 and 9 is conducted across all 26 districts
Classes	Classes 2 to 7 for the census self-evaluation assessment
	Classes 3, 5, 7 and 9 for the sample diagnostic assessment
Achievements	Gujarati, Maths, EVS, Science and Technology, Social Science, Hindi/English
Indicators	More than 100 parameters are adjudged, including academic proficiency, co-curricular activities, availability and utilization of infrastructure etc.
	Decided by Gujarat Council of Educational Research and Training, members of the State Resource Group of Sarva Shiksha Abhiyan for various subjects, DIET principals and representatives from the Department of Primary Education

Table 13:
The Role of Each Type of Assessment in Gunotsav

Gunotsav Assessments	The Role They Play
Self Assessment (SA)	Messaging importance of learning outcomes and their accountability to teacher and larger educational community
	Provides teacher an understanding of the achievement levels in his/her classroom
	Covers all teachers and students
Officer Assessment (OA)	Gives a signal of seriousness to education community
	Involvement of senior officers in education and understanding the key ground level issues
Diagnostic Assessment (DA)	Actionable feedback of learning gaps, common errors, misconceptions, strong and weak competencies
	Rigorous methods for tracking improvement annually
	Objective and controlled testing process using trained evaluators
	Full-length tests with question-wise feedback
	Representative sample provides rigour at low cost and effort (1/10th of size of OA)

Source: Table reproduced from the Learning Curve – Sankar 2013

Instruments

Subject matter experts at the Gujarat Council of Educational Research and Training (GCERT) prepare the test instrument for self and official evaluation. EI supplies the items in these assessments that are inference and critical thinking based.

Test Administration

Selection of schools

All primary schools take part in the self-evaluation component of Gunotsav each year. Further, 25% of schools in each block are selected at random to participate in the state-official led assessment later in the year. Visits are not pre-determined and schools are informed of their participation only one day ahead of the assessment. Lastly, a representative sample of approximately 1,100 schools takes the diagnostic learning assessment. The test administration broadly follows the calendar as specified in Table 14.

School self assessment

The school self-evaluation is conducted at the end of the first quarter of the academic year, around September or October.

The principal is encouraged to conduct the self-evaluation personally. The evaluation is across various learning and non-learning indicators and the methodology to conduct each is provided in Gunotsav guidelines by the GCERT. The school grades itself on a 0-10 scale for each indicator. The indicators that Gunotsav assesses are specified in Table 15.[46]

Table 14:
Gunotsav 2011 Annual Calendar

Period	Activity	By Whom/What Level
Jul-Aug	Informing District Education Depts and BRCCs of the two phases of assessment in schools	Dept of Education
Aug-Sep	Designing self-evaluation booklets and officer evaluation booklets and first phase-assessment papers for Classes 3 to 4	GCERT/DIET/State SRG members
Oct	Self-evaluation by all government primary schools	All government primary schools
	Data entry of self-evaluation	At the cluster level
	Making list of schools for each government officer's visit	District level
Nov	Schools take up in-class remedial activity/ Extra classes for children	
Dec	Briefing and understanding the process	
	Evaluation by senior government officials	25% of schools in the state
	Data entry of evaluation done by senior government officials	
Jan	Sharing of Gunotsav results	State level
Jun-Jul	Sharing of Gunotsav results	SSA and DPEO

Source: Table reproduced from Gunotsav GCERT

Table 15:
Domains Evaluated through Gunotsav

Academic Activities	Co-curricular Activities	Utilization of School Infrastructure
Reading, writing and computing skills	Prayer and yoga	Sanitation facility
Workbooks	Cultural activities	Drinking water facility
Science practical book and map book	School library	Electrical facility
Science and Maths	Exhibitions, use of computer lab	Cleanliness
Time of school and local visits	Sports meets	School health
Attendance	VEC, MTA and PTA meetings	School garden maintenance and cleanliness

Source: Table reproduced from Gunotsav GCERT

Exhibit 6:
Self-evaluation Form to Indicate Teacher Training Needs

Seeking Suggestions for Training Needs/Methodologies

Till date you have received manifold trainings. With the help of these trainings, your teaching skills must have improved a lot. There is scope for improvement in many subjects and subject details. There is also a possibility of bringing about many positive changes in the training methods.

If these points are duly considered beforehand for further training programmes, then the quality can definitely be improved with a scientific approach. With this purpose in mind a suggestion list has been prepared. You can suggest important points for your training needs through this list.

It is very important to assess yourself in this process. You must keep in mind that your training will be organized according to suggestions.

Instruction Guidelines:-

1) The suggestion list given to you has 4 sections.
   a. Training for teaching skills
   b. Training for subject matter skills
   c. Training for professional skills
   d. Your choice of training method
2) Give your preference for the kind of training you require. Grade the training method of your choice from no. 1 in decreasing order
3) Same number of preference can be given to one or more training methods

Teacher's name: _____ School Name: _____

Source: Gunotsav, GCERT www.gunotsav.org/gunotsavaheval.aspx

Official assessment

Two months after the self-evaluation, government officials conduct an officer evaluation. The Department of Primary Education informs all Class I and Class II officers in Gujarat and the evaluation is preceded by several advertisements and media attention. This is conducted to create a sense of urgency and high priority for the event. The officers receive a short briefing from the Chief Minister either in Gandhinagar, or through videoconference in all the district offices. 3,000 Class I and II officers visit close to 9,000 schools over a period of three days. One or two department members and a local liaison officer accompany each official.

The team of officials spends an entire day in the selected school. The team arrives at the school before the day begins and starts by participating in the prayer session. This is followed by academic and non-academic evaluation, engaging with students during the mid-day meal and reviewing school infrastructure. At the end of the school day, the officials interact with parents and understand their views on education-based activities in the community. The School Management Committee also participates in this process, as it gives them a chance to voice their views with a larger audience and makes them feel more accountable in their role. The evening is spent in the community, as students present a cultural show and when officers return, they upload their data on the Gunotsav school evaluation website.

The structure of each subject's officer evaluation is detailed below:

- **Reading –** for each class-level, the officer is provided with 100 to 150 paragraphs of reading tasks. 20% of the students in the class are randomly selected to read one paragraph. Each student is given 2-3 minutes to read and then graded on the basis of their competency.
- **Writing –** officers choose a writing task from their tool-kit and teachers transact it under their guidance. The task is then evaluated by the teachers and some responses are then reviewed by the officer team.
- **Maths –** This is similar to the writing assessment and officers fill marks in teacher and student evaluation sheets.

Student learning assessment

Student learning assessment of all students from Classes 2 to 7 or 8 is conducted, on the academic curriculum of the grade below their current level. Evaluation of reading and writing in Gujarati language is conducted first. Following this, Maths, English, Social Science and Environmental Science are tested. 100 to 150 items are prepared to test the competencies and then distributed at random.

After the academic testing concludes, teachers check and grade students by class and by subject. This is essential, as the self-evaluation form requires that teachers indicate which classes and subjects they have taught in the current and previous year. This data is utilized to provide grading for teachers and students.

Sample diagnostic assessment

A sample detailed, diagnostic assessment is conducted at the end of Gunotsav to understand how well children in the state are learning. This assessment has a different

objective to the self and official evaluation and is conducted with approximately 1 lakh students in 1,100 schools across Classes 3, 5, 7 and 9. The test provides detailed information regarding the strengths and weaknesses in student learning. To ensure that the information is scaffolded, EI conducts both 'written' and 'group oral' papers. The focus of the items is to ensure that students are 'learning with understanding', as shown in the sample question in Exhibit 7.

EI conducts a series of capacity building workshops for state and district personnel to develop the latest skills in building and using student assessments. Additionally, post data analysis, dissemination workshops are held for teachers to understand the insights from the data and incorporate the information in classroom practices. This includes discussion video responses of students, to understand their misconceptions.

Reports and Analysis

The Gunotsav GCERT team prepares compact discs with district data for each of the 26 districts. The data includes grading of teachers, schools, learning achievement of students, CRC Coordinator (CRCC) and the cluster, BRC Coordinator (BRCC) and the block. It also includes a district report at a glance. Additionally, SSA distributes printed certificates for each of the above.

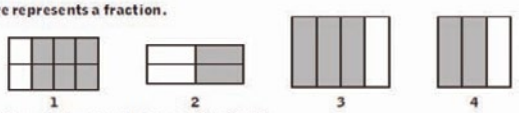
At the culmination of Gunotsav, each district receives the following reports and data sets:

- A grade summary of all the schools, per block. This includes a summary sheet and details of each school-cluster code, cluster name, village code, village name, school code, name of school, average grade of academic indicators, average grade of other parameters and total grade. This report card provides a snapshot of the number of schools in each block that have a ranking from A – F.
- A block evaluation sheet (for all the blocks within the district), which includes the name of the BRCC, evaluation details of all schools and the final grade of the block.
- A cluster evaluation sheet (for all the clusters within each block), which includes the name of the CRCC, evaluation details of all schools and the final grade of the cluster.
- A school evaluation sheet which includes the name of the school, the village, the cluster and all the teachers, with the final evaluation details and grade of the school.
- A teacher report card for each individual teacher that includes the teacher's identity code, name of school, learning outcomes achieved by students who were taught by the teacher and the final grade appointed to the teacher.

Use of Results

Schools in Gujarat are graded annually on the basis of their performance in Gunotsav. The evaluation gives 70% weightage for academic performance and 30% for school infrastructure and other parameters. Report cards are prepared for 33,450 schools and 1,72,000 teachers. A grade summary is available for districts, blocks, schools and teachers. Block certificates for BRCCs and cluster certificates for grading CRCCs are also given based

Exhibit 7:
Sample Gunotsav Diagnostic Assessment Item

S.No	Traditional format	Alternative forms testing the same concept – Testing for 'Learning with Understanding'
1.	What is the reduced form of 6/9?	Each figure represents a fraction.  1 2 3 4 Which two figures represent the same fraction? A. 1 and 3 B. 1 and 4 C. 2 and 3 D. 3 and 4
		1b. Write a fraction that is larger than 2/7
2.	Add: 7.234 + 21.34	2a. Which of these numbers is CLOSEST to 423.1? A.4231 B. 4.23 C.42.3 D.423
		2b. Which of these numbers is the largest? A. 7.234 B. 6.1 C. .4999 D. 21.34

Source: Exhibit reproduced from The Learning Curve – Sankar 2013

on the performance of the schools. Schools that only conduct self-evaluation in a certain year are assessed solely on these grades.

However, in schools where an officer-evaluation has been conducted, the following evaluation process is followed:

- If the officer-evaluation grade is higher than the self-evaluation grade, an average of both grades is considered final
- If the officer-evaluation grade is lower than the self-evaluation grade, the officer-evaluation grade is considered final
- If the officer-evaluation is less than 1.5 times the school self-evaluation, the final evaluation is half of the officer evaluation

From this school evaluation, several school improvement programmes have been initiated. SSA has helped conduct remedial programmes for 3 months post the officer assessment. A separate time for language and arithmetic remediation has been allotted in the school calendar. Schools have been encouraged to create additional Teaching Learning Material (TLMs) and supplementary grants have also been provided for this purpose. Students are also provided with take-home workbooks for summer vacations.

As mentioned earlier, all district officers in Gujarat have videoconferencing facilities available. Teacher training institutes, such as DIETs, have utilized these, to showcase student centred learning activities. Civil-society organisations have partnered with the GCERT in this process in some specific geographic areas. Pratham, for example, conducted remedial programmes in 11 districts.

Furthermore, 'Chintan Baithaks' (reflection sessions) are organized at the block level to facilitate the sharing of data and classroom experiences amongst teachers and principals. In these sessions, teachers form cluster groups of 10-12 to discuss student performance, socialize strategies that have worked and set targets for improvement. Similarly, school principals gather to make vision plans, which encompass learning and non-learning indicators and discuss their progress in achieving these.

CASE STUDY 3: MADHYA PRADESH

The Mandate

Since the introduction of the Right to Education Act (RTE) in 2009, Madhya Pradesh has strived to make quality education available to every child. Pratibha Parv is an annual assessment initiative in this quality agenda, managed and administered by the Rajya Shiksha Kendra (RSK).

History

Pratibha Parv was launched in the academic year 2011-2012. It is a comprehensive and holistic assessment programme to assess and evaluate the achievement level of students, while also monitoring school activities and infrastructure.

All government primary and upper-primary schools in the state – approximately 1.12 lakh schools, with more than 100 lakh enrolled students - are assessed on the same day twice in every academic year.

Purpose

The stated objectives of the Pratibha Parv are:

- To assess the academic performance of students and track it at regular intervals at the elementary education level
- To create improvement in the quality of education in primary and upper-primary schools
- To bring awareness about quality education amongst teachers and the general public
- To assess overall infrastructure available in the school
- To verify the availability and use of school facilities
- To provide a social audit opportunity and develop a sense of educational ownership in society

Assessment at a Glance

Table 16:

Pratibha Parv Snapshot

Parameter	Description
Structure	All students across primary and upper-primary classes in all government schools in Madhya Pradesh
Scale	1,12,788 schools across the state, covering more than 3,50,000 teachers and 1,00,00,000 students
Frequency	Biannually – at the end of Semester 1 and Semester 2
Classes	1, 2, 3, 4, 5, 6, 7, 8
Achievements	For Classes 1 – 4: English, Hindi, Maths
	For Classes 5 – 8: English, Hindi, Maths, Sanskrit (or other second language), Science, Social Science
Indicators	Learning indicators and also 20+ school system indicators

Instruments

Pratibha Parv comprises of a learning assessment and a non-learning survey that assesses school infrastructure, facilities and daily activities.

The learning assessment is designed by the RSK's Monitoring & Evaluation Department, with the support from the Curriculum team, UNICEF and the Azim Premji Foundation. Teachers from leading private schools in the state also provide feedback. The examination is largely knowledge-based and assess minimum level competencies aligned to the state curriculum.

The instrument allocates 10 marks for each subject in the examination and all subject questions are collated into one paper, that the students take over a two-hour duration. That is, students in Classes 1 to 4 answer a 30-mark paper (10 for English, Hindi and Math, respectively) and students in Classes 5 to 8 answer a 60-mark paper (10 for English, Hindi, Math, Sanskrit, Science and Social Science respectively). This restricts the number of question items in the instrument. Furthermore, Class 1 to 4 test carry five marks for

oral evaluation and five marks for dictation. Class 5 – 8 students also have to answer some written questions.

The non-learning indicators that are assessed during Pratibha Parv include:

- Daily prayer proceedings
- Students' uniform, hygiene and cleanliness
- Availability of clean drinking water
- The use of school radio facilities
- The use of school computer facilities
- Physical education classes
- Mid-day meal provision, including the kitchen shed, utensils, quality of food and distribution procedure
- Availability and use of TLMs
- School library, including the availability of books, inventory and a distribution register
- Student textbooks, exercise books, use of the classroom
- Formation, activation and participation of the School Management Committee
- School building, including availability and physical condition of rooms and notice-boards
- Pupil-teacher ratio and teaching quality
- Total student enrolment, attendance and measures taken to bring out-of-school children back
- Achievement of disadvantaged and differently-abled students

The instrument has undergone some changes through the three annual cycles. Most notably, students in Classes 5 – 8 now have to answer some written questions, reflecting the need to improve written communication skills. Earlier, their test comprised solely of multiple choice questions. Class 1 and 2, especially in Maths, are now pictorial rather than being largely text based. Lastly, the 2013 cycle of the Pratibha Parv carried three different sets for each question paper, so that a broader list of test items could be utilized.

Interestingly, moving away from global practice, participants tick answers on the test and teachers mark these later. This is because students have found it difficult to answer on OMR sheets.

Test Administration

Selection of schools
All primary and upper-primary government schools in the state participate in Pratibha Parv biannually.

School self-assessment
The self-assessment component of the Pratibha Parv is limited to analysis of non-learning indicators. School principals are given this list of non-learning indicators and conduct a self-review of their school's performance. For each indicator, they can select whether their performance is 'satisfactory' or 'unsatisfactory'. An external assessor then corroborates this review and a final score is allotted.

Peer team assessment

The Pratibha Parv external assessment takes place on two specific days annually. These days are scheduled in advance.

For the first assessment of the year, the Pratibha Parv team at the RSK conducts training for 50 DIET principals and 50 academic staff, who then conduct workshops at the district and block level. Block Resource Coordinators (BRCs) then conduct short, structured sessions with two groups of people:

1. **Secondary and higher secondary teachers –** These teachers are seconded to primary and upper primary schools for the first Pratibha Parv assessment of the year that is conducted in December or January. The training lasts approximately one hour and describes the responsibilities of these teachers. Their key role is to monitor the student learning assessment and conduct an independent survey of the non-learning aspects of the school.

2. **School principals –** The training with school principals is two hours long and describes the form for self-review of non-learning indicators.

The process of external administration for the first Pratibha Parv is standardized throughout the state. The RSK sends a circular in advance to the school with the testing dates. On these dates, students of both morning and afternoon school shifts arrive in the school at 9.30 a.m. The school principal, in the presence of the external administrator, an SMC member and two students, opens the assessment question papers that are sent to the school the day before in sealed packages.

The teachers conduct the assessment in student classrooms over the course of two hours. The external administrator monitors all the classes during this time. The question-paper is hand-marked by teachers the same afternoon, under the guidance of the external administrator and student achievement data is input into common scoring sheets, which record item-level responses. These data-recording sheets are checked for quality assurance by the external administrator and then sent to the nodal BRC within a day. The BRC then inputs the results to the Madhya Pradesh Education Portal directly over a fortnight.

From 2013, the RSK has endeavored to second external administrators, i.e. secondary or higher secondary teachers to schools within two kilometers of their regular teaching location. This makes travel easier for the teachers and also reduces costs associated with daily allowances.

The second assessment is usually conducted in March/April, before the students' last summative examination of the academic year. It follows a similar pattern, except a state official, usually a Class I or Class II officer from the Government of Madhya Pradesh conducts it. State officials are allotted schools to monitor across the state.

Reports and Analysis

The RSK generates several reports, on learning and non-learning indicators, based on assessment data. Most of these are available online for public review. The available reports include:

- A summary of infrastructure and teacher learning materials available in schools, including future requirements
- A summary of teacher training needs for the state, on the basis of low performing sections of the assessment
- A list of school ratings – A, B, C, D or E – on the basis of performance on both Pratibha Parv assessments and the final summative assessment of the year
- Teacher and student grading – per classroom, school, cluster, block and district, also on a similar A to E scale

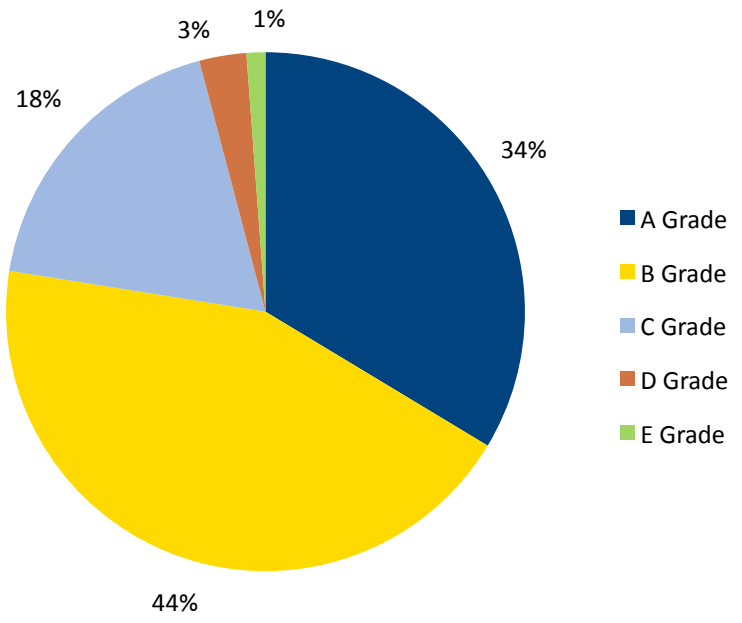
The general public can use longitudinal data available on the Education Portal to also compile reports.

Use of Results

The RSK allots special supporting measures to schools and students who receive D or E grades. For students, schools develop a 'School Action Plan' to specifically coach and mentor them. The school also shares these results with parents at progress meetings. This has helped engage parents in the school community to ensure student retention and attendance.

Exhibit 8 below showcases school performance in the 2012 Pratibha Parv.

Exhibit 8:
Pratibha Parv School Report



Total schools	A Grade	B Grade	C Grade	D Grade	E Grade
1,12,779	37,932	49,490	20,724	3,308	1,325

Source: Graph reproduced from Centre for Innovation in Public Systems

Furthermore, the RSK ranks districts on the basis of their school grades and this compilation is shared with the Chief Minister. The RSK presents district specific results with each District Collector, who then discusses it with all the block level officials. These results feed into the Annual Work Plan that each district prepares for the next academic year.

Lastly, the Pratibha Parv rewards commendable assessment performance. Schools that have all classes in the Grade A zone (for Pratibha Parv 1 and 2 and the last summative assessment of the year) are awarded ₹10,000 and each teacher with three classes in the Grade A zone is awarded ₹15,000 or ₹5,000 per class.

CASE STUDY 4: BRAZIL

The Mandate

The National Institute for Educational Studies and Research (INEP) is a federal agency under the Ministry of Education responsible for national assessment and evaluation of education, including making diagnoses and recommendations.

History

Brazil began to give attention to standardized assessment in the late 1980s. Both the 1988 Constitution and the 1996 Law of Directives and Bases of National Education stress the importance of assessing the educational system. The Evaluation System of Basic Education (SAEB) was introduced in 1991 and underwent significant methodological innovations in 1995. Prova Brasil was first implemented in 2005.

Purpose

Both SAEB and Prova Brasil are used to assess the education system rather than individual students. The assessments' purpose is to enable educational authorities to make more informed decisions. While SAEB is a diagnostic instrument for the system as a whole, Prova Brasil assesses individual schools and municipalities with the aim of helping the government decide how to allocate technical and financial resources. Prova Brasil also

increases parental pressure on low-performing schools.

Assessment at a Glance

Table 17:

SAEB and Prova Brasil Snapshot

Parameter	Description
Scale	SAEB is administered to a sample of public and private schools. Prova Brasil is administered to all public schools with at least 20 children in the class assessed. For schools that participate in both SAEB and Prova Brasil, the two tests are effectively one as INEP disaggregates data from Prova Brasil to get a sub-sample for SAEB
Frequency	SAEB and Prova Brasil are administered every two years
Classes	SAEB assesses Classes 4, 8 and 11. Prova Brasil assesses Classes 4 and 8 only
Achievements	SAEB and Prova Brasil assess student learning in Portuguese language and Maths. In Portuguese the focus is on abilities in reading and in math the focus is abilities related to problem solving, ideas of space and shape, numbers and operations, measurements and information interpretation

Instruments

Socioeconomic questionnaires are administered along with the exams to collect information that would be associated with student performance.

Reports and Analysis

SAEB results are reported by state and Prova Brasil provides data at the level of schools and municipalities.

Use of Results

The results of Prova Brasil for public schools in Classes 4 and 8 and SAEB for private schools in Classes 4, 8 and 11 and public schools in Class 11 are used for Brazil's Basic Education Development Index (IDEB), a measure of educational quality. A school's IDEB score is calculated as a multiple of performance on Prova Brasil and the promotion rate, which is used to ensure that students are not held back or encouraged to drop out in order to improve the school's score. Scores are calculated on a 1 to 10 scale, which is aligned with PISA scores.

IDEB scores are provided at the level of school, municipality and state for public schools in Classes 4 and 8 and only at the state level for Class 11 in public schools, where school-level data is not available.

The IDEB score is used to set individual goals for each school. They are only provided at the state level for private schools to achieve average PISA performance in 2021. Targets are set on a two-year basis and schools, municipalities and states are expected to create plans for meeting these targets.

For low-performing schools, the state must create improvement plans, including the technical and financial resources needed from the Ministry of Education. The Ministry then allocates additional resources to low-performing schools and municipalities, such as teacher training programmes provided by federal universities.

Individual states also make decisions about how to use IDEB results. For example, the state of Ceará worked with the state's federal university to train and certify professionals to work with teachers and principals in their improvement efforts. Ceará also has an incentive programme for the 150 highest-performing schools and an assistance programme for the 150 lowest-performing schools. The highest performers receive additional funding for their school and have a mandate to partner with a low-performing school and help it improve, while the lowest performers receive additional training, instructional resources and other assistance from the state.

Other Points of Interest

Additionally, INEP has developed Provinha Brasil, an early grade assessment of reading, which is intended to be used by teachers and school directors to assess students' learning levels and respond to difficulties they may have. It is administered to students at the beginning and end of Class 2. INEP also conducts an annual school census to gather data at the school, municipal and state level on the number of students and teachers and on school infrastructure.

CASE STUDY 5: AUSTRALIA

The Mandate

The Australian Curriculum, Assessment and Reporting Authority (ACARA) was established by a 2008 Act of the Australian Federal Parliament. The Standing Council on School Education and Early Childhood, a forum for the coordination of national policy on education amongst state, territory and federal ministers, provides direction to ACARA. ACARA is responsible for:

- National curriculum from Foundation to Year 12, including content and achievement standards
- A national assessment programme aligned to the national curriculum
- Data collecting and reporting that supports research, analysis, resource allocation and accountability

The National Assessments Program – Literacy and Numeracy (NAPLAN) is the most significant component of the overall assessment programme.

History

Prior to the introduction of NAPLAN, each state had its own assessment for literacy and numeracy. The implementation of NAPLAN in 2008 was part of a broader context of

educational reform. The Council of Australian Government, an intergovernmental forum for national and state/territory governments, approved the National Education Agreement (NEA) in 2008 as part of its overall National Productivity Agenda. The NEA focuses on outcomes, including that "Young people are meeting basic literacy and numeracy standards and overall levels of literacy and numeracy achievement are improving," and includes a significant reporting framework.

Purpose

In 2008, Australian Education Ministers jointly established the Melbourne Declaration on Educational Goals for Young Australians, which are that

- Australian schooling promotes equity and excellence
- All Australians becomes successful learners, confident and creative individuals and active and informed citizens

In the Melbourne Declaration Commitment to Action, Ministers agreed on eight actions to achieve these goals, including "promoting world-class curriculum and assessment."

Assessment at a Glance

Table 18:
NAPLAN Snapshot

Parameter	Description
Scale	NAPLAN is a census assessment. Students with a language background other than English, who arrived in Australia less than a year before the tests and students with significant disabilities may be exempted from testing. Parents or other caregivers may also withdraw students from NAPLAN on the basis of religious or philosophical objections
Frequency	Annual
Classes	3, 5, 7 and 9
Achievements	Reading, writing, language conventions (spelling, grammar and punctuation) and numeracy

Instruments

NAPLAN is made up of four tests for Classes 3 and 5, one in each of the domains assessed. For Classes 7 and 9, the numeracy assessment is made up of two tests, one where calculators are allowed and one where they are prohibited. The questions on the reading, language convention and numeracy assessments are either multiple choice or require a short written response. For the writing test, students respond to a persuasive writing prompt. The same writing prompt is used for all years.

Test Development

The test development process is as follows:

1. ACARA reviews and revises guidelines for test development
2. ACARA contracts out question development to outside organisations and test

developers from these organisations write questions

3. Test managers from each state and territory and NGO representatives review questions to make sure they meet curriculum and jurisdiction-based circumstances

4. Representative samples of students from each state and territory complete trial assessments and the data is analysed to determine which of the questions meet required specifications

5. ACARA's Expert Advisory Panel, made up of five measurement and assessment experts, reviews the resulting tests as well as the trial data and makes additional suggestions

Test Administration

In each jurisdiction, a designated Test Administration Authority is responsible for NAPLAN administration in accordance with the National Protocols for Test Administration. NAPLAN is administered over three days in May, one each for language conventions and writing, reading and numeracy. The length of testing in each domain ranges from 40 to 80 minutes.

ACARA aims to transition to online delivery of the NAPLAN beginning in 2016.

Reports and Analysis

ACARA contracts out central data analysis to a third-party agency, which analyzes the raw data and carries out an equating process so that NAPLAN tests in different years can be reported on the same scale. They provide performance results to the states and the Centre, which then design and disseminate reports.

Student performance on each domain of NAPLAN is evaluated on a scale from 1 to 10, with each band in the scale reflecting increasing sophistication of skills. Six bands of the scale are used to report student achievement at each class level and one band is demarcated as the national minimum standard for that year. E.g. the Class 3 report shows bands 1 to 6 and band 2 is the national minimum standard, while the Class 5 report shows bands 3 to 8, with band 4 as the national minimum standard. Reporting student performance at all years across this same scale allows for monitoring student progress over time. An equating process is used so that results from NAPLAN tests in different years can be reported on the same scale.

Data is reported at the level of individual students, schools and states/territories. Reports on individual students are provided to families and schools. These reports show student performance in each domain in relation to the national minimum standard, the national average, the performance of the middle 60% of students in that class level and the school average in some states/territories. These reports also include a summary of what skills students have typically demonstrated at each band level for each test. Parents generally receive these reports in September.

ACARA also publishes an annual national report, which provides data at the national and state territory level on participation and performance for each domain in each class level.

The report also shows performance by gender, indigenous status, language background for students whose parents speak a language other than English at home, parental occupation, parental education and location, metropolitan, provincial, remote and very remote. A sample of this report is included in Appendix M. A summary of the report is published in September and the full report is published in December. States also publish their own reports.

ACARA publishes school performance on its *My School* website, where users can also compare results of one school against other schools with students from what it calls "statistically similar backgrounds". ACARA uses an Index of Community Socio-educational Advantage, a measure of demographic factors shown to influence students' educational outcomes, to determine which schools should be considered similar.

Use of Results

The results of NAPLAN are used to monitor performance, promote accountability and inform policy making at both national and jurisdiction levels. In Queensland, for example, relatively low performance on NAPLAN encouraged reform and data on the performance gap for Indigenous students stimulated the development of new programmes. At the national level, NAPLAN performance informs federal funding. For example, from 2008-2012 the federal government allocated up to AUD 540 million (₹2976 crore) to support programmes that improved literacy and numeracy. Over the first two years, money was allocated to each state and territory based on its share of students at or below national minimum standards in literacy and numeracy. Over the last two years, money was allocated to states based on their success in meeting their specific improvement targets, as reflected through NAPLAN results in conjunction with other indicators. States also used NAPLAN results to determine which schools were eligible for participation in this programme.

Schools also get significant feedback on student performance, which helps to promote the use of NAPLAN results at the school level. Results, however, are reliable only at the level of the five overall achievements assessed: reading, writing, spelling, grammar and punctuation and numeracy and so are not intended to on their own provide more specific diagnoses for individual students.

Challenges

In their review of evaluation and assessment in Australia, Santiago, Donaldson, Herman and Shewbridge (2011) documented stakeholders concerns about an overemphasis on NAPLAN, particularly due to its primacy on the *My School* website. Stakeholders raised concerns about the potential negative impact on low-performing schools and their students, who might be labeled as failures, as well as perverse incentives for schools to rig their results, such as by encouraging low-performing students to stay home on test day. The media has also used NAPLAN results to publish league tables ranking schools, which stakeholders saw as inappropriate and misleading. One suggestion the reviewers made was that the Australian government consider expanding the scope of information on the *My School* website to include school evaluation reports.[47]

The review noted that there are risks that the emphasis on NAPLAN could have a restrictive impact on classroom teaching and learning, if the curriculum is narrowed based

on the basic skills covered in the current tests. Although ACARA discourages excessive preparation, observers have noted that there is a significant phenomenon of "teaching to the test" and putting lots of time into preparation. This phenomenon also undermines the validity of results. There are also concerns about NAPLAN undermining the centrality of teacher-based assessment.[48]

Santiago et. al. suggested that NAPLAN may have cultural biases that make it an unfair test for Indigenous students, despite considerable efforts to make the test inclusive. More data is needed on how the test functions for disabled students.

The OECD reviewers also found that Australian policy has emphasized the accountability function of assessment rather than the improvement function. They suggested that there could be greater focus on defining a national vision for how the existing data, including the results of the NAPLAN test, can be used to bring about improvement in school and classroom practices and that teachers would benefit from additional training in how to use NAPLAN data.

Other Points of Interest

The National Assessments Program also includes sample-based assessments in science literacy, civics and citizenship and information and computer technology (ICT) literacy. Each assessment is administered once every three years to students in Classes 6 and 10 (Class 6 only for science literacy). Australia also participates in The Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS).

CASE STUDY 6: CHILE

The Mandate

The Agencia de Calidad de la Educaciõn (Education Quality Assurance Agency), established by law in 2011, is responsible for evaluating student learning outcomes and other indicators of educational quality. It also classifies schools based on performance, provides improvement guidelines to schools and informs the community about school performance.

Prior to the creation of the Agencia de Calidad de la Educaciõn, Chile's national assessment programme – Sistema de Medición de la Calidad de la Educaión (SIMCE) – was first administered by an external agency, the Pontificia Universidad Católica de Chile and then by the Ministry of Education.

History

SIMCE, which began in 1988, was the first national assessment in Latin America. The assessment was originally created to help parents make decisions about school choice in the context of a national voucher policy. At the time of creation, the assessment was also intended to drive improvement by promoting competition between schools, provide information needed to evaluate educational policies and enable pedagogical

improvements. The assessment was created under a right-wing dictatorship and then substantially revised following the return to democratic rule in 1990.

Purpose

The purpose of SIMCE is to improve quality and equity in education. SIMCE is intended to contribute to improvement through three main levers: informing policy, providing pedagogical support to educators and holding schools accountable.

Assessment at a Glance

Table 19:

SIMCE Snapshot

Parameter	Description
Scale	SIMCE is a census exam in all subjects except information and computer technology, physical education and special needs. Also, very small schools inaccessible locations are excluded
Frequency	Annual (though subjects other than reading, writing and Maths are administered less frequently)
Classes	2, 4, 6, 8, 10, 11 (Class 11 is assessed only on English and is the only class assessed in this subject)
Achievements	Maths, Natural Sciences, Social Sciences, English, Information and Computer Technology, Physical Education, reading, writing and special needs

Instruments

The tests are mostly in multiple-choice format, but open-ended questions and essays have been gradually introduced. Along with the assessments, students, parents and teachers respond to questionnaires, which include questions on socioeconomic status and teaching qualifications.

Test Development

Teams from the Agencia de Calidad de la Educaciõn and external test developers trained by the agency develop test items. Selection panels review the items and those the panels approve are trialed in schools in conjunction with the census assessment.

Test Administration

Private contractors are responsible for the assessment field operations. Test administration is standardized across the country.

Reports and Analysis

SIMCE results are calculated at the school, regional, school and student level and information is widely disseminated in a variety of formats as explained in detail by Ramirez (2012). Staff of the Agencia de Calidad de la Educaciõn does data analysis and an equating process is used so that results are comparable across years.

The SIMCE National Report provides national and regional mean scores for each

assessment; the percent of students at each performance level: beginner, intermediate and advanced; mean scores by socioeconomic background, gender and public/private school; and trends in mean scores across years.

A report is also provided to each school, which shows the national and school mean scores, as well as comparisons of school performance against last year's performance, the national mean and the mean of other schools serving students from the same socioeconomic background. Schools reports also include the percentage of students by performance level, analysis of the content and skills required to answer sample test questions and guidelines for implementing workshops to enable the use of assessment results.

National results and school-level performance are also published on the SIMCE website and in a newspaper supplement. The newspapers generally print a ranking of schools when the SIMCE data is released.

Parents receive reports on their child and school's performance, as well as recommendations to support student learning. These reports are intended both to promote accountability for the school and to involve parents in their child's educational process. There is also a geo-referential tool for parents, which shows schools and their mean scores on Google Maps.

For researchers, there are school and student data files and data analysis tools. To obtain access to student data, researchers must commit to not using the results to identify students or teachers. Ramirez (2012)'s complete explanation of the SIMCE dissemination strategy is available in Appendix N.

Use of Results

According to Ramirez (2012), SIMCE has served to centre attention on student learning. Results are used first and foremost to inform policymaking. The information is used to monitor quality and equity and to design and evaluate intervention programmes. Providing additional resources to the lowest-performing schools is a key aspect of intervention. Under the P-900 programme, SIMCE identifies the 900 schools that are the lowest performers on the Language and Maths tests. These schools then receive support with infrastructural improvements, textbooks and books for classroom libraries, teaching materials and in-service workshops for teachers.

SIMCE results are also used to promote greater school accountability through the widespread publication of results in the Chilean media and through incentive programmes. For instance, the National Performance Assessment System (Sistema Nacional de la Educación or SNED) programme uses SIMCE scores, in conjunction with four other measures of educational quality, to award monetary incentives to teachers from the best-performing schools. In addition, the Preferential Subsidy programme (Subvención Escolar Preferencial – SEP) provides financial incentives and pedagogical assistance to schools serving low-income students that meet agreed targets. Additionally, the Quality Assurance Law of 2011 allows for closing schools that do not show improvements.

Results are used to a lesser extent by schools and teachers to make pedagogical decisions at the school or classroom level. One reason for this may be that educators lack training in how to use the information, especially given that assessment literacy receives little

attention in teacher education programmes. The Ministry of Education tried to remedy this problem with seminars and workshops but many schools were unable to fully implement these trainings. Ramirez (2012) notes that a 2011 law, which requires schools to create improvement plans based on SIMCE results and other indicators, has good potential to strengthen the use of SIMCE results at the school level.

Challenges

SIMCE lacks an established mechanism to monitor how the assessment results are used and the overall impacts of the assessments programme. Also, as mentioned above, the assessment is currently not used within schools themselves to the extent envisioned.

Ramirez (2012) explains that there are also concerns about possible unintended negative consequences associated with SIMCE. The publication of results in newspapers, for example, may lead to stigmatization of the lowest-performing schools and the poorest students, a point of particular note given that 80% of the variance in school mean scores can be attributed to socioeconomic status. Other potential issues are teachers teaching to the test, by focusing on those subject areas and by over-using multiple choice questions in the classroom; schools disproportionately directing resources to the tested classes; and the further segregation of the school system by academic performance and socioeconomic background.[49]

Section Seven

Appendices

APPENDIX A

Sample National Assessment Project Plan

Main activity and sub activities	Duration	Effort (hours)	Start date	Finish date	Person
Plan and convene steering committee meeting.	**1 month**	**40**			
Identify and contact participants.					
Determine suitable date for meeting.					
Organize transportation, venue, accommodation, meeting, and refreshments.					
Send out invitations.					
Specify an assessment framework.	**1 month**	**120**			
Select sample of schools.	**2 months**	**160**			
Specify target population.					
Contact Department of Education for school data.					
Prepare school and within-school sampling procedures.					
Draw sample.					
Finalize sample.					
Develop instruments	**4 months**	**640**			
Develop, edit, and finalize items and scoring guides.					
Identify item writers.					
Appoint item writers.					
Train item writers.					
Draft test items, sample items and administration manual.					
Review test items.					
Pilot test items.					
Develop scoring guides.					
Score test items.					
After formal review, select final set of test items and sample items.					
Complete artwork and test layout.					
Estimate time allowed for each test.					
Prepare administration manual and scoring guides.					

Source: Howie, S. (2004). A national assessment in mathematics within an international comparative assessment: research article. Perspectives in Education, 22 (2).

APPENDIX B

Brief on TIMSS and PIRLS

About TIMSS and PIRLS

For the past 20 years, TIMSS (Trends in International Mathematics and Science Study) has measured trends in mathematics and science achievement at the fourth and eighth grades. It has been conducted on a regular 4-year cycle since 1995, making TIMSS 2011 the fifth assessment of mathematics and science achievement trends. TIMSS Advanced, which measures trends in advanced mathematics and physics for students in their final year of secondary school, was conducted in 1995 and 2008, and is scheduled for 2015 (with the sixth assessment of TIMSS). For the past 15 years PIRLS (Progress in International Reading Literacy Study) has measured trends in reading comprehension at the fourth grade. First assessed in 2001, PIRLS has been on a regular 5-year cycle since then. Most recently, PIRLS was expanded in 2011 to include prePIRLS, which is a less difficult version of PIRLS. Both TIMSS and PIRLS were assessed in 2011, when the cycles of both studies came into alignment.

In general, participating countries use TIMSS and PIRLS in various ways to explore educational issues, including: monitoring system-level achievement trends in a global context, establishing achievement goals and standards for educational improvement, stimulating curriculum reform, improving teaching and learning through research and analysis of the data, conducting related studies (e.g. monitoring equity or assessing students in additional grades), and training researchers and teachers in assessment and evaluation.

TIMSS and PIRLS results are disseminated through reports and via the web through a well-documented international database for within and across country research.

APPENDIX C

Brief on the PISA

Brief

The Programme for International Student Assessment (PISA) is an international assessment of the skills and knowledge of 15-year-olds. PISA assesses students' performance on 'real-life' tasks that are considered relevant for effective participation in adult society and for life-long learning.

PISA is implemented every three years, starting in 2000, and 2012 saw the fifth implementation of the study. The number of countries participating in PISA has increased from 32 in 2000 to 64 in 2012, making it the largest study of its kind.

The subject areas or 'domains' assessed by PISA are reading, mathematics and science. In each cycle of PISA, one of these subjects is the main focus ('major domain') of the assessment, with less emphasis placed on the remaining subjects ('minor domains').

Occasionally, additional domains are assessed. For example, problem solving was a minor domain in 2012.

PISA is gradually transitioning to a computer-based assessment, so that by 2015, most countries will be administering PISA entirely by computer. In 2012, as well as completing paper-based assessments of mathematics, reading and science, students in over 40 countries including Ireland completed computer-based assessments of mathematics, reading, and cross-curricular problem solving.

Along with assessing the achievements of students, PISA collects detailed contextual information from students, parents, and principals through the context questionnaires. In Ireland, a national teacher questionnaire is implemented in each cycle.

PISA is a project of the Organisation for Economic Co-operation and Development (OECD).

Participants	PISA assesses 15- and 16-year-old students, since in most OECD countries students at this age are approaching the end of compulsory schooling. First, a representative sample of schools is selected (reflecting a mixture of schools by size, location, type, and gender and socioeconomic composition). Next, students are sampled within participating schools. Up to 35 students are selected at random in each school. In schools with fewer than 35 students in the eligible age range, all such students are selected. In Ireland, participating students are in both junior cycle and senior cycle. About 60% of students are in third year, 25% in transition year, and 15% or so are in fifth year. PISA has rigorous technical standards, so response rates must be high. At least 85% of schools, and 80% of students, need to participate, in order for results to be deemed valid. Teachers who are selected to participate in the national teacher survey are teachers who teach the 'major domain'. For example, in PISA 2012, all mathematics teachers in each PISA school were invited to participate. In PISA 2015, for the first time, parents of PISA students will be invited to complete a parent questionnaire. The number of countries/regions participating in PISA has increased from 32 in 2000 to 73 expected to participate in PISA 2015. Since 2003, all OECD member countries have taken part in PISA.

Source: Table reproduced from Educational Research Centre, Ireland - www.erc.ie/?p=55

APPENDIX D

Brief on the EGRA

Early Grade Reading Assessment (EGRA)

Why early grade reading? The ability to read and understand a simple text is one of the most fundamental skills a child can learn. Without basic literacy there is little chance that a child can escape the intergenerational cycle of poverty. Yet in many countries, students enrolled in school for as many as six years are unable to read and understand a simple text. Recent evidence indicates that learning to read both *early* and at a sufficient *rate* are essential for learning to read well. Acquiring literacy becomes more difficult as students grow older; children who do not learn to read in the first few grades are more likely to repeat and eventually drop out, while the gap between early readers and nonreaders increases over time.

What is the Early Grade Reading Assessment? Most national and international assessments are paper-and-pencil tests administered to students in grades four and above (that is, they assume students can read and write). Results for those few low-income countries that participate in PISA or TIMSS (and inferring from the results of regional assessments such as PASEC and SACMEQ)[1] indicate that the median child in a low-income country performs at about the third percentile of a high-income country distribution. From these results we can tell what students did not know, but cannot ascertain what they *did* know (often because they scored so poorly that the test could not distinguish whether the child did not know the content or simply could not read the test).

In contrast, EGRA is designed to orally assess the *most basic foundation* skills for literacy acquisition in early grades, including pre-reading skills such as listening comprehension. The test components are based on recommendations made by an international panel of reading and testing experts and include timed, 1-minute assessments of letter naming, nonsense and familiar words, and paragraph reading. Additional (untimed) segments include comprehension, relationship to print, and dictation. In each of the language pilots conducted to date, EGRA meets psychometric standards as a reliable and valid measure of early reading skills.

How is EGRA administered? EGRA is an individually administered oral assessment of foundation literacy skills requiring about 15 minutes per child. It has been designed as an inexpensive and simple diagnostic of individual student progress in reading. In addition, ministry personnel can use the results to identify schools with particular needs and develop instructional approaches for improving foundation skills (e.g., poor letter naming results may indicate the need for additional alphabet practice).

Where has EGRA been used and what do the results look like so far? EGRA has been applied in over forty countries and in a host of languages. Results thus far indicate generally low levels of student acquisition of foundation literacy skills. For example, in one country, students at the end of grade 2 were correctly able to name about 23 letters in English in one minute. For the sake of comparison, in the United States, reading norms state that children reading at fewer than 40 correct letters per minute at the end of kindergarten should be considered at some risk, while those reading at less than 27 are definitely at risk. Thus, the *average* level of letter reading fluency in the tested country was, in grade 2, *half* of what in the United States would be considered to put the child at some risk at the end of kindergarten.

To provide an overall sense of levels of reading in the countries where EGRA has been tried, the following table provides summary averages for oral reading fluency in terms of correct words per minute (cwpm). Note that country names have been excluded to avoid comparisons (cross-language comparisons are not encouraged due to differences in language structure; see below). In Africa's English-speaking countries, grade 2 oral reading fluency is 10 to 20 words per minute. In the United States, students are expected to read about 50 cwpm at the end of grade 1—thus, EGRA countries are performing below the fifth percentile level of U.S. norms. With a relatively transparent language

[1] Organisation for Economic Co-Operation and Development's Program for International Student Assessment (PISA); Trends in International Mathematics and Science Study (TIMSS); Programme d'Analyse des Systems Educatifs de la Confemen (PASEC); Southern Africa Consortium for the Measurement of Educational Quality (SACMEQ).

Oral Reading Fluency Levels (Correct Words per Minute) in EGRA

		Grade		
		1	2	3
Africa	French	2.9	17.4	32.4
(Low Income)	English 1	2.2	4.0	9.2
	English 2		11.4	
Latin America	English		59.0	73.1
(Lower Middle Income)	Spanish 1	9.2	29.3	
	Spanish 2	32.0	59.6	78.8

such as Spanish, performance standards are higher: Children in Spain achieve about 60 cwpm at the end of grade 1—EGRA countries are performing at one half that rate. Oral reading fluency is both an excellent predictor of later reading skills (correlations between 0.7 and 0.9, using high-income country studies) and a warning light: If reading problems are not corrected early on, the gap in reading skills between readers and nonreaders actually increases. Thus, a key task in low-income countries is to get all children reading well by the end of grade 1, or at the latest by grade 2 where scripts are complicated and poverty is rampant.

Can EGRA results be used to compare results across languages and countries? Preparation of the EGRA instrument for use in a particular country generally involves some adaptation, including translation into the language of instruction. It is important to recognize that this limits the ability to make comparisons across countries. One reason for this stems from the research on reading acquisition, which indicates that while all children move through the same stages when learning to read, the rate at which they move through them differs by language (and the degree to which these languages vary in their orthographic complexity). Another reason is related to the technical standards for making such comparisons, which require evidence that translation and other adaptations do not change the difficulty level of the test and hence the meaning of the scores across systems. Despite the challenge of comparing results across countries and languages, finding out at which grade children are typically "breaking through" to literacy, and comparing these grades across countries or regions, will be a useful analytical and policy exercise.

Once EGRA identifies the areas for improvement, what can be done to improve learning outcomes? EdData II has developed a strategy for improving student learning using research-based instructional approaches to remedy critical areas identified by the EGRA instrument. For example, based on the EGRA results, teachers may be taught to monitor students' oral reading fluency and practice decoding strategies. This approach recently was tested in 40 randomly selected schools (20 treatment and 20 control) in Kenya, and a two-year control-treatment intervention is under way in Liberia. This continuous cycle for improving student learning, including evaluation together with specific support for teachers and monitoring for accountability, is a process that has generated average student learning gains on the order of 30 percent or more in South Africa (District Development Support Program) and Zambia (Break Through to Literacy). Efforts in Mali and Niger using EGRA to inform the development of materials and sequenced, scripted teaching and continuous assessment strategies have demonstrated very promising results, even for large classrooms (all children reading within a few months). Research in the United States indicates that early acquisition of foundation literacy skills is an important predictor of later school success; teachers can promote that success by strengthening those skills identified as needing improvement by the EGRA instrument.

How can USAID Missions "buy into" EdData II and how much will it cost? RTI International holds the current EdData II task order contract from USAID/Washington. EdData is a USAID-funded program that has supplied survey-based data on education in countries worldwide since 1997. The data are used for planning, monitoring, and evaluating education policies and programs. Missions can issue a Request for Task Order Proposal (RFTOP) as part of EdData II, detailing the proposed number of schools and students to be evaluated and the degree of representation required for sampling. Costs for application of the EGRA instrument will vary by country and are highly dependent on local inputs of labor and transportation. In countries where ministry staff or other salaried officials are trained as enumerators, the cost of application can be significantly reduced. As a purely illustrative example, local assessment costs (excluding international technical assistance) in The Gambia in 2007 for a baseline testing of 1200 students in 400 schools was about $25,000 (included workshop costs, per diems, printing, and transportation for Ministry staff; paid enumerators were not used for this task). Additional information, supporting documentation and references, and sample instruments can be found at www.eddataglobal.org.

APPENDIX E

Brief on the NAS

An introduction to National Achievement Survey

The National Achievement Surveys (NAS) are important to know the health of the education system and to provide feedback to the policy planners, researchers and practitioners. NAS serves a unique role within the education system of India in measuring and monitoring standards of learning achievement. Under SSA, Department of Education Measurement and Evaluation, NCERT has carried out two cycles of NAS for Classes III, V and VIII in the last decade.

Under NAS the same assessment instruments are administered in all participating states and union territories. This allows each state to compare its results with those of other states as well as with national aggregates. It also allows progress over time to be evaluated. NAS reports information on student performance for the nation and the states in selected subject areas. NAS also presents findings for different groups including results by gender, school location and social status. Unlike the examinations conducted by the various states and all-India boards, NAS does not produce results for individual students. Rather benchmarks are established based on different levels of performance.

The third cycle of NAS for Class V has been significantly changed to incorporate international best practices in student assessment systems. Steps have been taken to improve all aspects of achievement survey i.e. sampling, quality of tests and questionnaires, analysis and reporting. The Item Response Theory (IRT) which is used in all international surveys has been used for test construction and analysis to report the findings for Class-V NAS.

What is IRT and what are its advantages?

In previous national surveys, learning achievement data was analysed using Classical Test Theory (CTT) and average scores were reported as the 'proportion of answers correct'. This approach, whilst valid, has significant limitations. Most importantly, the results are more valid to particular tests and groups of students tested. The generalisation to other groups and linking to other similar tests is problematic under this system. This precludes use of multiple tests or to link results from one cycle to another. In order to overcome this limitation, current NAS cycles have used Item Response Theory (IRT), to link multiple test booklets and to analyse the data. This is in keeping with the best practice of major international surveys such as Trends in International Mathematics and Science Studies (TIMSS) and Programme for International Student Assessment (PISA) and Progress in Reading Literacy Study (PIRLS).

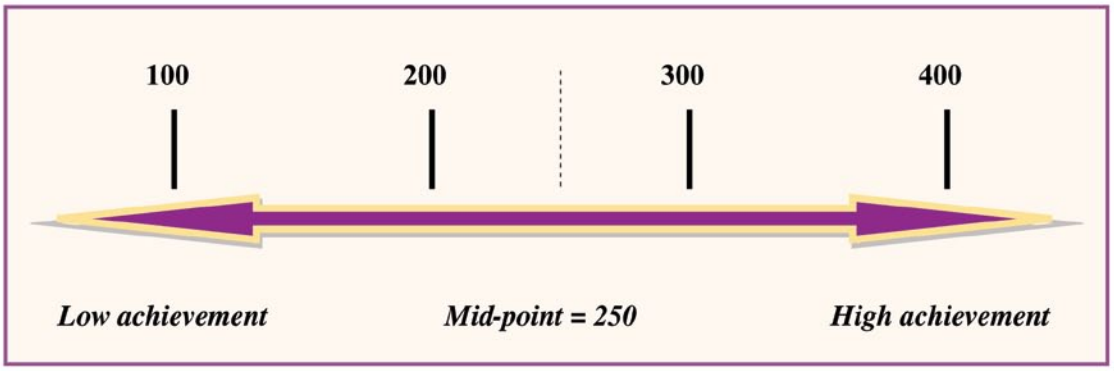
IRT uses a mathematical model to link a student's chance of answering correctly a particular item to two main factors: the student's level of ability and the item's level of difficulty. Whilst this method

makes the analysis more complex than traditional methods, it has many advantages.

- IRT places students and test items on the same numerical scale. This enables us to produce meaningful 'maps' of item difficulty and student ability.
- In IRT, the difficulty parameter for an item does not depend on the group of test takers. This allows us to use multiple test booklets to increase measurement points in any subject and these can be linked.
- IRT also allows us to compare scores from tests used in different cycles - an essential characteristic for monitoring progress over time.

Scale Scores

IRT is applied to calculate scale scores which are different from 'proportion of answers correct.' The score given to each item is not 0 or 1 but it depends upon the difficulty level of item and the range of scale used. The scale is fixed so that results from future surveys can be reported on the same scale. So if, for example, over a three-year period a state's average score in Mathematics rises from 248 to 254, we can compare these directly and draw meaningful conclusions about changes in student achievement. This is possible because, even though the scores have been derived from different students taking different tests at different times, the reporting scale is fixed.



In NAS class V, the scale chosen is from 0 to 500. The average score for the whole population is initially set at 250. However, if educational standards improve the overall average will rise from this 'baseline'. The standard deviation of the scale is initially set at 50 for the whole population. This means that the majority of students (about 70%) will have scores in the range 200 to 300. On this scale, a score of more than 400 would represent an extraordinarily high level of achievement (see figure below).

Interpretation of Scale Scores in NAS Class-V

The average score is reported for each participating state and UT. These are accurate for the sample chosen, but the true average for the population may vary from the sample average. The 'Standard Error' is an estimate of the likely variation. As a rule of thumb, the average score of the population will fall in

a range of plus or minus two standard errors from the sample average. For example, in the table below, the average Mathematics score of all the states listed is '251' and the sampling error is estimated to be '0.7'. This means that we can be confident that the true average for this group of states is in the range 251 ± 1.4.

When comparing two average scores, the standard errors of each must be taken into account. For example, in the table, State X has a mean score of '257' which looks higher than the group average score of 251. However, when we take the standard errors into account we see that the difference between this state's performance and that of the group is not statistically significant. Similarly, State X has a higher mean score than State P, but the observed difference in Maths achievement in these two states is not statistically different when the standard errors are considered. Both are 'average' states when it comes to Maths achievement.

Mathematics			
State or UT	Average Score	Standard Error	Significant Difference
State P	252	2.6	●
....	....	....	....
....	....	....	....
State X	257	3.2	●
State Y	298	3.1	◠
State Z	241	2.7	◡
Group Average	251	0.7	

● The state's average score is not significantly different to that of the group

◠ The state's average score is significantly above that of the group

◡ The state's average score is significantly below that of the group

Percentile Scores

In addition to average scores, NAS reports 'percentile scores'. Percentile tables and figures in NAS report illustrate the achievement within states at different percentiles. A percentile score indicates the scale score below which a certain proportion of students fall. For example, the 10th percentile score means that 10% of students may be found at or below it. (Hence, 90% of students can be found above it.) As shown in the exemplar table below, NAS reports list the scores achieved by students at key percentiles. Among these are the 25th (first quartile), 50th (second quartile or median), and 75th (third quartile) percentile. The range between the 25th and 75th percentiles (the inter-quartile range) represents the performance of the middle 50% of students. Hence, this is a good indicator of the state's degree of homogeneity in terms of the achievement of its students.

State or UT	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile	Range 75-25	Range 90-10
....	....	....	....	....	....	....	....
State 1	185	212	228	271	291	59	107
State 2	178	204	230	275	321	71	143
State 3	185	212	226	248	273	36	88
....	....	....	....	....	....	....	....

In the table above, States 1 and 3 have similar median scores (228 and 226 respectively). However, State 1 has a significantly higher score at the 75th percentile than State 3 (271 compared with 248). This shows that whilst the average scores for the two states are comparable, the top 25% of students in State 1 are doing significantly better than their peers in State 3. By providing such data, NAS allows States to compare achievement not only for 'average students', but also across the full ability range.

Target Population and Scope of NAS Class V

NAS is designed to assess learning achievement of students in the government system at the level of the state or union territory and thereby also for the entire nation. Hence, the target population was all children studying in government and government-aided schools in Class V across the country. Representative samples of students in all states are treated in the same way in order to provide a common and stable measure of achievement.

Sample schools included those managed by the Department of Education, Tribal/Social Welfare Departments, and Local Bodies as well as Private-but-government-aided schools. The survey collected data from 1, 22,543 students, and 10,851 teachers from 6,602 schools across 27 states and 4 UTs.

Three tests in each of the three subjects i.e. Language, Mathematics and EVS were administered in the sampled schools. Three questionnaires were also administered to capture background data of students, teachers as well as schools in which the students were studying.

Class V NAS Report

The Class V NAS report is an aggregated national level report that starts with an introductory chapter besides an Executive Summary. Chapter numbers 2-7 of the report focuses on the learning achievement in three subjects namely; Mathematics, Language and EVS each followed by the description of "what class-V students know and can do" in respective subjects. The last three chapters are based on different background variables derived from the questionnaires that influence student learning. The report is followed by appendices at the end and these provide insight into sampling design, IRT model and state level data table on key variables.

Source: National Council of Educational Research and Training. (2012). National Achievement Survey Class V. New Delhi: Author.

APPENDIX F

NAS Supporting Questionnaires

National Achievement Survey

CLASS 3

TEACHER QUESTIONNAIRE

To be filled in by the Field Investigator

❶ Name of the State _____		State Code ☐☐
❷ Name of the District _____		District Code ☐☐
❸ Name and Address of the School _____		School Code ☐☐☐☐
❹ Name of the Teacher _____	(Put a tick [✓] mark)	Subject(s) teaching: Lang. 1 Maths. 2 Lang. and Maths 3
❺ Sex	(Put a tick [✓] mark)	Sex: Male 1 Female 2
❻ Category	(Put a tick [✓] mark)	Category: SC 1 ST 2 OBC 3 Others 4
❼ Highest educational qualifications	(Put a tick [✓] mark)	Upper Primary 1 Sec. 2 Sr. Secondary 3 Graduation 4 Post-graduation and above 5
❽ Language as subject studied up to (*The Language in which sampled students are being tested*)	(Put a tick [✓] mark)	Upper Primary 1 Sec. 2 Sr. Secondary 3 Graduation 4 Post-graduation and above 5
❾ Mathematics as subject studied up to	(Put a tick [✓] mark)	Upper Primary 1 Sec. 2 Sr. Secondary 3 Graduation 4 Post-graduation and above 5
❿ Professional Qualification	(Put a tick [✓] mark)	Untrained (No Certificate / Diploma/Degree) 1 Graduate Level (B.Ed., LT etc.) 3 Elementary Teacher Training Certificate/Diploma/JBT/BTC 2 Post-graduate (M.Ed.) 4
Date of Survey D D M M Y Y Y Y		
Name of the Field Investigator _____		Signature _____
Data Scrutinized by _____		Signature _____ Date _____

Instructions for filling up the Teacher Questionnaire (*To be filled in by the Field Investigator only*)

- ⇨ *All entries are mandatory and should be in **English only**. Write all codes in international numerals, i.e.,* **1, 2, 3...**
- ⇨ This questionnaire is for the purpose of collecting information about the teacher.
- ⇨ It has no bearing on individual headmasters or teachers.
- ⇨ Information provided, will be used only for a National Study to know the health of the educational system.
- ⇨ Tick [✓] in the appropriate box against the question for giving response.
- ⇨ At the most 2 teachers are to be taken who are teaching **Language** and **Mathematics** subjects to the sampled students of Class III.
- ⇨ A separate teacher questionnaire is to be filled for both the teachers.

विद्यया ऽमृतमश्नुते
एन सी ई आर टी NCERT
Educational Survey Division
NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING
Sri Aurobindo Marg, New Delhi

2012-13

Put a tick [✓] mark in the appropriate box for registering response against each question

TEACHING EXPERIENCE AND TRAINING

11 Employment Status

1	2	3
Regular full time	Adhoc/Temporary/ Against leave vacancy	Para teacher/ Shiksha karmi/ etc.

12 Total teaching experience in Primary Classes

1	2	3	4	5	6
0-1	1-3	4-6	7-10	11-15	15+

13 In-service training program attended in the session of 2012-13

1	2
Yes	No

14 Training Programme attended during 2011-12

Organizer		No. of Prog.		No. of days	
TRC	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
Sch. Comp.	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
CRC	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
BRC	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
DIET	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
SCERT *	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
SSA	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
NCERT	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	
OTHER	1 Yes, 2 No	A 1, B 2-3, C 4-5, D 5+		A 1, B 2-3, C 4-5, D 5+	

Note: Tick the SCERT * if any one of the following is functional in your state– **DSERT, DTERT, GCERT, MSCERT, SIERT, SIE, Directorate of School Education**

TEACHING AND EVALUATION PRACTICES

15 Availability and Use of Teaching Aids
(Tick only in one box in one row.)

	Not Available	Regularly	Some-times	Never
i. Teacher's Handbook	1	2	3	4
ii. Charts	1	2	3	4
iii. Maps	1	2	3	4
iv. Globe	1	2	3	4
v. Mathematics Kit	1	2	3	4
vi. Self prepared TLM	1	2	3	4
vii. TLM from other sources	1	2	3	4
viii. Books other than textbooks	1	2	3	4

16 Homework to students (in a week's time)

1	2	3	4
Never	Once	2-3 times	More than 2 to 3 times

17 Do you maintain Teacher's Diary?

1	2
Yes	No

18 How many periods do you have per week?

1	2	3	4	5
upto 24	25-30	31-35	36-40	41 or More

19 Do you use revised textbooks based on NCF-2005 for Class III ?

1	2
Yes	No

20 Have you attended any training programme based on NCF-2005 ?

1	2
Yes	No

21 Types of Examination / Evaluation school have?

	1 Yes	2 No
i. Observations		
ii. Activity based		
iii. Oral		
iv. Unit / Monthly test		
v. Term test		
vi. Half yearly test		
vii. Annual examination		

22 In your school, how severe is each problem?
(Tick only in one box in one row.)

	Not a Problem	Minor Problem	Serious Problem
i. Repairing of the school building	1	2	3
ii. Large and more classrooms	1	2	3
iii. More workspace for teachers	1	2	3
iv. Materials for experiments/activities	1	2	3

23 How would you characterize each of the following within your school? *(Tick only in one box in one row.)*

	Low	Medium	High	Not Sure
i. Teachers' job satisfaction	1	2	3	4
ii. Teachers' degree of success in implementing the curriculum	1	2	3	4
iii. Teachers' expectations for student achievement	1	2	3	4
iv. Parental support for student achievement	1	2	3	4
v. Parental involvement in school activities	1	2	3	4
vi. Students' desire to do well in school	1	2	3	4

National Achievement Survey

CLASS 3

SCHOOL QUESTIONNAIRE

To be filled in by the Field Investigator

❶ Name of the State _____		State Code ☐☐
❷ Name of the District _____		District Code ☐☐
❸ Name and Address of the School _____		School Code ☐☐☐☐
❹ Location of the School	(Put a tick [✓] mark)	Rural 1 Urban 2
❺ Type of School	(Put a tick [✓] mark)	Boys 1 Girls 2 Co-Ed. 3
❻ Indicate by ticking [✓] in the appropriate box up to what class the school is?		1 (1-4) 2 (1-5) 3 (1-7) 4 (1-8) 5 (1-10) 6 (1-12) Others _____ 7
❼ Is Pre-primary section attached to your school?	(Put a tick [✓] mark)	1 Yes 2 No
❽ School managed by	(Put a tick [✓] mark)	State Govt. / Zila Parishad / Panchayat / Local body / Municipal Committee 1 Govt. aided 2
Date of Survey D D M M Y Y Y Y		
Name of the Field Investigator _____		Signature _____
Date of Data Scrutiny D D M M Y Y Y Y		
Name of the District Coordinator _____		Signature _____

Instructions for filling up the School Questionnaire

- ⇨ *This questionnaire is to be filled in by the Field Investigator only.*
- ⇨ *All entries are mandatory and should be in **English language only**. Write all codes in international numerals, i.e.,* **1, 2, 3...**
- ⇨ This questionnaire is for the purpose of collecting information about the school.
- ⇨ It has no bearing on individual schools, headmasters or teachers.
- ⇨ It is to be used for a National Study to know the health of the educational system of the country.
- ⇨ Provide correct information about the school.
- ⇨ Tick [✓] in the appropriate box against the question for giving response.

विद्यया ऽ मृतमश्नुते
एन सी ई आर टी NCERT

Educational Survey Division
NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING
Sri Aurobindo Marg, New Delhi

2012-13

Put a tick [✓] mark in the appropriate box for registering response against each question

⑨ TEACHERS : Number of teachers teaching Primary Classes

		1	2	3	4	5	6
		None	1	2-3	4-6	7-10	11 or More
A. Regular full time	Male	1	2	3	4	5	6
	Female	1	2	3	4	5	6
B. Adhoc /Temporary/ Against leave vacancy	Male	1	2	3	4	5	6
	Female	1	2	3	4	5	6
C. Para teacher/ Shiksha karmi/ etc.	Male	1	2	3	4	5	6
	Female	1	2	3	4	5	6

⑩ Whether the school has got School Grants for the year (2012-13) under SSA? 1 Yes 2 No

⑪ Number of days school worked for the academic year (2011-12)

1	2	3	4	5	6
upto 160	161-180	181-200	201-220	221-240	241 or more

⑫ Duration of a period (in minutes)

1	2	3	4	5
25	30	35	40	45+

⑬ Number of periods per day

1	2	3	4	5
5	6	7	8	9+

⑭ Number of working days per week 1 5 day 2 6 day

⑮ SCHOOL FACILITIES : State about availability of the following.

A. Teaching aids

i. Maps 1 Yes 2 No
ii. Globe 1 Yes 2 No
iii. Charts 1 Yes 2 No
iv. Maths Kit 1 Yes 2 No
v. Science Kit 1 Yes 2 No
vi. Library 1 Yes 2 No

B. Ancillary facilities

i. Musical Instruments 1 Yes 2 No
ii. Annual medical check up for children 1 Yes 2 No
iii. First-aid kit 1 Yes 2 No

C. Physical facilities

i. Sports and games material 1 Yes 2 No
ii. Safe drinking water 1 Yes 2 No
iii. Toilet facilities 1 Yes 2 No
iv. Separate toilet facilities for girls 1 Yes 2 No
v. Electric connection for the school 1 Yes 2 No
vi. Playground 1 Yes 2 No
vii. Mats and furniture for students 1 For All 2 Some 3 None
viii. Television 1 Yes 2 No
ix. Computer 1 Yes 2 No
x. Telephone connection 1 Yes 2 No
xi. Staff room 1 Yes 2 No
xii. Type of building 1 Pakka 2 Partial Pakka 3 Kachcha
xiii. Number of classrooms for Primary sections

1	2	3	4	5	6
None	1-2	3-4	5-6	7-8	9 or More

PARTICIPATION

⑯ How many times BRC / CRC personnel visited the school in the academic session 2011-12?

1	2	3	4	5	6	7	8	9
None	1	2	3	4	5	6-7	8-9	10 or More

⑰ Does school have the following

PTA 1 Yes 2 No
MTA 1 Yes 2 No
VEC 1 Yes 2 No
SMC 1 Yes 2 No
AEC 1 Yes 2 No

⑱ Which plays most important role in the functioning of the school? *Tick any one of the following:*

1 VEC/SMC/AEC 2 MTA 3 PTA

PTA- *Parent Teacher Association*, **MTA**- *Mother Teacher Association*, **VEC**- *Village Association Committee*, **SMC**- *School Management Committee*, **AEC**- *Area Education Committee*

CLASS 3

National Achievement Survey

PUPIL QUESTIONNAIRE

2012 - 13

विद्यया ऽमृतमश्नुते

एन सी ई आर टी NCERT

Educational Survey Division

NATIONAL COUNCIL OF EDUCATIONAL RESEARCH AND TRAINING

Sri Aurobindo Marg, New Delhi

Name of the State _____ ☐☐

Name of the District _____ ☐☐

Name of the School _____ ☐☐☐☐

Location of the School Rural (1) ☐ Urban (2) ☐

Student ID	Name of the Student	Q1 Gender	Q2 Age	Q3 Category	Q4 Belongs to BPL family	Q5 Belongs to Physically Challenged
01		☐	☐☐	☐	☐	☐
02		☐	☐☐	☐	☐	☐
03		☐	☐☐	☐	☐	☐
04		☐	☐☐	☐	☐	☐
05		☐	☐☐	☐	☐	☐
06		☐	☐☐	☐	☐	☐
07		☐	☐☐	☐	☐	☐
08		☐	☐☐	☐	☐	☐
09		☐	☐☐	☐	☐	☐
10		☐	☐☐	☐	☐	☐
11		☐	☐☐	☐	☐	☐
12		☐	☐☐	☐	☐	☐
13		☐	☐☐	☐	☐	☐
14		☐	☐☐	☐	☐	☐
15		☐	☐☐	☐	☐	☐
16		☐	☐☐	☐	☐	☐
17		☐	☐☐	☐	☐	☐
18		☐	☐☐	☐	☐	☐
19		☐	☐☐	☐	☐	☐
20		☐	☐☐	☐	☐	☐

Pupil Questionnaire to be filled in by Field Investigator only

General Instructions:

- All entries are mandatory and should be in **English language only**. Write all codes in international numerals, i.e., **1, 2, 3...**
- This questionnaire is for the purpose of collecting information about the student background.
- It has no bearing on individual student.
- It is to be used for a National Study to know the health of the educational system of the country.
- Please provide correct information about the student.
- Fill the appropriate code in the box given in the question column for giving response.
- Copy the **Student ID** from **Field Notes** where you allotted Student ID to sampled students.
- Ensure that **Student ID** written here must be the same as in **Field Notes** as well as mentioned at **Language** and **Mathematics tests** to an individual student.
- Complete all entries on this page from **School Register** by taking the help of Class Teacher/Principal/Head Teacher for all sampled students.

Use Codes for questions as follows:

Q1. Gender: Use '**1**' for Boy, '**2**' for Girl

Q2. Age: Write in completed years i.e. if a child is of 7 years 5 months and 27 days then write 07 and if a child is 7 years 6 months and 1 day then write 0 8

Q3. Category: Use '**1**' for SC, '**2**' for ST, '**3**' for OBC and '**4**' for Others

Q4. BPL: If a student belongs to Below Poverty Line (BPL) family then write '**1**' if not then write '**2**'

Q5. Physically Challenged: If a student belongs to physically challenged group write '**1**' if not then write '**2**'

Student ID	Q6			Q7		
	Educational Status			Occupational Status		
	Father	Mother	Guardian	Father	Mother	Guardian
01	☐	☐	☐	☐	☐	☐
02	☐	☐	☐	☐	☐	☐
03	☐	☐	☐	☐	☐	☐
04	☐	☐	☐	☐	☐	☐
05	☐	☐	☐	☐	☐	☐
06	☐	☐	☐	☐	☐	☐
07	☐	☐	☐	☐	☐	☐
08	☐	☐	☐	☐	☐	☐
09	☐	☐	☐	☐	☐	☐
10	☐	☐	☐	☐	☐	☐
11	☐	☐	☐	☐	☐	☐
12	☐	☐	☐	☐	☐	☐
13	☐	☐	☐	☐	☐	☐
14	☐	☐	☐	☐	☐	☐
15	☐	☐	☐	☐	☐	☐
16	☐	☐	☐	☐	☐	☐
17	☐	☐	☐	☐	☐	☐
18	☐	☐	☐	☐	☐	☐
19	☐	☐	☐	☐	☐	☐
20	☐	☐	☐	☐	☐	☐

Pupil Questionnaire to be filled in by Field Investigator only

Complete all entries on this page from **School Register** by taking the help of Class Teacher/Principal/Head Teacher for all sampled students.

Use Codes for questions as follows:

Q6. Educational status: Information will be provided in Guardian column, only if student is living with a Guardian rather than his/her Father/Mother.

0 - Father is not alive or Mother is not alive or both are not alive

1 - Illiterate (not able to read and write)

2 - Literate (no formal schooling but can read and/or write)

3 - Education upto Primary level

4 - Education upto Elementary level/Middle (Class VII/VIII)

5 - Education upto Secondary level (Class X)

6 - Education upto Higher/Senior Secondary level (Class XII)

7 - Education upto Degree and above (Academic or Professional or Both)

8 - Information not available

Q7. Occupational status : Information will be provided in Guardian column, only if student is living with a Guardian rather than his/her Father/Mother.

0 - Father is not alive or Mother is not alive or both are not alive

1 - If parent/guardian is Unemployed

2 - Household / Housewife

3 - Agricultural Labour/Domestic Servant/ Daily Wager / Street Vender

4 - Farmer (Cultivator)

5 - Skilled Worker / Office Worker

6 - Shopkeeper / Businessman

7 - Teacher / Lecturer / Professor

8 - Manager / Senior Officer / Professional

9 - Information not available

Student ID	Q8		Q9	Q10	Q11	Q12	Q13	Q14	Pupil Questionnaire to be filled in by Field Investigator only
	Number of		Attended Pre-primary Classes	Who Helps in study at home	Lang. used at home is same as in school	Take Private Tuition	Homework is given regularly	Homework is checked regularly	
	Brother	Sister							
01	☐	☐	☐	☐	☐	☐	☐	☐	
02	☐	☐	☐	☐	☐	☐	☐	☐	
03	☐	☐	☐	☐	☐	☐	☐	☐	
04	☐	☐	☐	☐	☐	☐	☐	☐	
05	☐	☐	☐	☐	☐	☐	☐	☐	
06	☐	☐	☐	☐	☐	☐	☐	☐	
07	☐	☐	☐	☐	☐	☐	☐	☐	
08	☐	☐	☐	☐	☐	☐	☐	☐	
09	☐	☐	☐	☐	☐	☐	☐	☐	
10	☐	☐	☐	☐	☐	☐	☐	☐	
11	☐	☐	☐	☐	☐	☐	☐	☐	
12	☐	☐	☐	☐	☐	☐	☐	☐	
13	☐	☐	☐	☐	☐	☐	☐	☐	
14	☐	☐	☐	☐	☐	☐	☐	☐	
15	☐	☐	☐	☐	☐	☐	☐	☐	
16	☐	☐	☐	☐	☐	☐	☐	☐	
17	☐	☐	☐	☐	☐	☐	☐	☐	
18	☐	☐	☐	☐	☐	☐	☐	☐	
19	☐	☐	☐	☐	☐	☐	☐	☐	
20	☐	☐	☐	☐	☐	☐	☐	☐	

General Instructions:

- Information for Question Number 8 to 14 are to be filled by the Field Investigator by **interview mode** i.e., calling each sampled student one by one in order to Student ID such as first 01, then 02, then 03 as so on...
- Use the information provided by the student for filling up of the entries in the Question Number 8 to 14.

Use Codes for questions as follows:

Q8. How many brother(s) and sister(s) you have?
0 - No brother/sister
1 to **9** - For student having actual number of brother(s)/sisters(s) separately

Q9. Have you attended Pre-primary classes?
1 - Yes **2** - No

Q10. Do you get help in study at home?
1 - Yes **2** - No

Q11. Is the language used in your home is the same as the language used in your lessons at school?
1 - Yes **2** - No

Q12. Do you take private tuition?
1 - Yes **2** - No

Q13. Do you get homework regularly from school?
1 - Yes **2** - No

Note : *If the answer of Q13 is 'No' then not to ask question number 14 and use code '0' for question number 14.*

Q14. Does your teacher check homework regularly?
1 - Yes **2** - No

Student ID	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Pupil Questionnaire to be filled in by Field Investigator only
	Problems in Maths are given	Problems in Maths are checked	Subject like the most	Like being in school	Borrowing books from School Library	Use Computer in School	Like to do the most	**General Instructions:** • Information for Question Number 14 to 20 are to be filled by the Field Investigator by **interview mode** i.e., calling each sampled student one by one in order to Student ID such as first 01, then 02, then 03 as so on.... • Use the information provided by the student for filling up of the entries in the Question Number 15 to 21. **Use Codes for questions as follows:** **Q15. Does your teacher give you problems to solve in Mathematics?** **1** - Yes **2** - No Note : *If the answer of Q15 is 'No' then not to ask question number 16 and use code '0' for question number 16.* **Q16. Does your teacher check problems solved by you in Mathematics?** **1** - Yes **2** - No **Q17. Which subject do you like the most?** **1** - Language **2** - Mathematics **3** - Other subject **Q18. Do you like to come to school?** **1** - Yes **2** - No Note: *If school does not have library then do not ask question number 19 and use code '0'* **Q19. Do you borrow books from your school library?** **1** - Yes **2** - No Note: *If school does not have computer then do not ask question number 20 and use code '0'* **Q20. Do you use Computer in the school ?** **1** - Yes **2** - No **Q21. Which activities of the following do you like the most?** **1** - Drawing / Painting / Sketching / Dancing / Music / Singing **2** - Playing sports / games **3** - Watching television and videos **4** - Reading textbook / comic / magazine / storybooks **5** - Playing on computer **0** - None / Can't say
01	☐	☐	☐	☐	☐	☐	☐	
02	☐	☐	☐	☐	☐	☐	☐	
03	☐	☐	☐	☐	☐	☐	☐	
04	☐	☐	☐	☐	☐	☐	☐	
05	☐	☐	☐	☐	☐	☐	☐	
06	☐	☐	☐	☐	☐	☐	☐	
07	☐	☐	☐	☐	☐	☐	☐	
08	☐	☐	☐	☐	☐	☐	☐	
09	☐	☐	☐	☐	☐	☐	☐	
10	☐	☐	☐	☐	☐	☐	☐	
11	☐	☐	☐	☐	☐	☐	☐	
12	☐	☐	☐	☐	☐	☐	☐	
13	☐	☐	☐	☐	☐	☐	☐	
14	☐	☐	☐	☐	☐	☐	☐	
15	☐	☐	☐	☐	☐	☐	☐	
16	☐	☐	☐	☐	☐	☐	☐	
17	☐	☐	☐	☐	☐	☐	☐	
18	☐	☐	☐	☐	☐	☐	☐	
19	☐	☐	☐	☐	☐	☐	☐	
20	☐	☐	☐	☐	☐	☐	☐	

Page 4

Source: National Council of Educational Research and Training, (2012), Quotation for Analysis of Data of National Achievement Survey.
Accessed at www.ncert.nic.in/announcements/tendors/pdf_files/NAS_C3.pdf

APPENDIX G

Snapshot of Quotation for NAS Data Analysis

QUOTATION FOR ANALYSIS OF DATA OF NATIONAL ACHIEVEMENT SURVEY AT THE END OF CLASS III

1. The National Council of Educational Research and Training (NCERT) is an apex organization set up by the Government of India, under the Ministry of Human Resource Development with Headquarters at New Delhi to provide academic and technical resource support to central and State Governments for school education.

2. In NCERT, the Educational Survey Division (ESD) has undertaken a nationwide 'Achievement Survey' of children at the end of the Class III of school education. The survey sample covers about 8,400 schools, 15,500 teachers and 1,35,000 students from 300 districts throughout the country. NCERT therefore invites quotations from interested agencies/ parties for analysis of the data of "National Achievement Surveys at the End of Class III".

3. In the 'National Achievement Surveys of Class III' the following data has been collected:
   i. For Class III, the Achievement Tests consist of two subject areas namely Language (25 items) and Mathematics (30 items) given in Table-1 as below:

Table-1

Sl. No.	Achievement Test	No. of Items per test	Length of record
Main Tests			
1.	Language	25*	44
2.	Mathematics	30	47

**Item no. 25 is subdivided in to three sub parts i.e. 25A, 25B & 25C.*

   ii. With a view to studying the influence of intervening variables such as home, school and teachers, data has also been collected in the Pupil, Teacher and School questionnaires as given in Table-2 as follows:

Table-2

Sl. No.	Questionnaire	Length of record
1.	School Information Sheet (SIS)	31 to 71 (Depends on Number of sections in the school)
2.	Pupil Questionnaire (PQ)	39
3.	School Questionnaire (SQ)	52
4.	Teacher Questionnaire (TQ)	75

   iii. Information on the approx. number of districts, schools, students, teachers and the number of Achievement Tests covered in National Achievement Survey at the End of Class III is given in Table-3 as follows:

Table-3A (Questionnaires)

No. of States/ UT	No. of Districts	No. of Schools	No. of Students	No. of Teachers
35	300	8,400	1,35,000	15,500

Table-3B (Response Sheets)

S. No.	Name of the Tool / Material	Quantity of material to be received from one school	No. of Items / Sub Items	Length of record
1.*	Language Response Sheet	1	25*	44
2.*	Mathematics Response Sheet	1	30	47
3.	School Information Sheet SIS	1	6 to 10	31 to 71 (Depends on Number of Sections in the school)
4.*	Pupil Questionnaire (PQ)	1	26	39
5.	School Questionnaire (SQ)	1	43	52
6.	Teacher Questionnaire (TQ)	Max. 2	52	75

** Each Language Response Sheet, Mathematics Response Sheet and Pupil Questionnaire (PQ) will contain responses of 20 students.*

4. The following tasks are required to be undertaken:
    a. **Transcription of data of Achievement Tests**: This will require to be undertaken as per the 'Framework for Plan of Data Analysis for National Achievement Survey Class III' attached at **Annexure 'A'**. It will be necessary to make convenient batches of response sheets, checking of identification codes and entry of information from response sheets to computer media with minimum **98%** accuracy in database. This transcription is required to be done for all States/ UTs. Sample copy of the response sheet is enclosed at **Annexure 'B'**.
    b. **Transcription of data from Pupil, Teacher and School Questionnaires:** This will also require to be made as per the 'Framework for Plan of Data Analysis for National Achievement Survey Class III' enclosed at **Annexure 'A'**. The Identification Code is on the cover page of each Questionnaire. Verification for 'no entry' in the Identification codes must be made and relevant entries carried out with corrected Identification Codes before data entry is started. As per the above-mentioned criteria, 98% of accuracy in data entry needs to be maintained universally. The sample questionnaires are enclosed at **Annexure 'C'**.
    c. **Transcription of Response Sheets:** This will be required for two Achievement Tests, namely Language and Mathematics for all the students in a School of a District in the States/ UTs.
    d. **Merging of existing data with common database for creation of separate databases:** This will be created by using Achievement Tests, Pupil Questionnaire (PQ), Teacher Questionnaire (TQ) and School Questionnaire (SQ).

e. **Aggregation of Achievement Scores:** This has to be done by aggregating the Achievement Scores with the student, school and teacher variables.

f. **Tabulation of data:** The 'Framework for Plan of Data Analysis for National Achievement Survey Class III' is attached at **Annexure 'A'.** This booklet contains dummy tables. These tables are to be generated by developing software in any language or by using any suitable package like SPSS or SAS. Outputs will be required in MS Excel format and MS Word. These tables are to be replicated for each state/UT and for the country.

g. **Multivariate Statistical Analysis of Data:** Regression analysis has to be carried out as per the procedure discussed in the subsequent paragraph:

- Merge the files containing test scores with the file of student records. Care should be taken to manage the missing or mismatched information in student record or test records.
- Existing variables will be recoded with directional and logical values.
- Principal Component Analysis will be carried out for giving weightage and scaling of variables by providing correlation matrix as an input.
- Variables need to be standardized before creating composite indices.
- Regression analysis separately for each state and country for Test Scores in Mathematics and Language as dependent variable and School, Teacher and Pupil variables will be treated as independent variables. Enter method needs to be used during regression analysis.

h. **Computation of co-efficient** of correlations and significance of differences within and between variables will be carried out by taking all the variables (Achievement, Pupil, Teacher, School)

i. **Competency/Area wise achievement** of students in Mathematics and Language has to be analysed so as to identify areas of learning difficulties.

J. **Item Analysis (Facility Value and Discrimination Index)** needs to be carried out item wise in Mathematics and Language to understand the nature of the test.

k. **Reliability co-efficient** of two tests (Mathematics and Language) will be carried out to know the consistency of the tests.

5. The 'Framework for Plan of Data Analysis for National Achievement Survey Class III' provides the format for Tables at **Annexure 'A'**.

6. The time frame for completion of the work is given in the Table-4 as below:

Table-4 (Time Frame in which the work is required to be completed)

Sl. No.	Work Description	States/ UTs	Time Schedule
1.	**Data Entry**/ Data Verification/Cleaning	States/UTs	23 days
2.	**Aggregation of Achievement Scores** with student variables, school variables and teacher variables by matching them on different parameters	States/ UTs and overall	5 days
3.	**Sample checking** by NCERT Faculty	5 % random check	3 days
4.	**Statistical Analysis of Data**	States/ UTs and overall	5 days

5.	**Tabulation of data and g**eneration of tables as per 'Plan of Data Analysis for National Achievement Survey Class III' i.e. Annexure 'A'	All 35 States/ UTs and for the country	3 days
6.	Printout of the tables	States/ UTs and overall	1 day
		Total Time	**40 days**

GENERAL TERMS AND CONDITIONS:

1. Interested firms are invited to submit their quotation for the tasks described in 1 – 6 as above. The quotations should be submitted in two parts, viz., **(a) Technical Quotation and (b) Financial Quotation**. These should be placed in two different envelopes to facilitate evaluation of Technical Quotations before the Financial Quotation is opened.
2. The **Technical Quotation** must be placed in an envelope super-scribed "**Technical Quotation**" and must contain the details specified in **Annexure 'D'** including the following:
    a. Details of available personnel, along with the Curriculum Vitae of the Lead Systems Analysts/ Data Base Administrators.
    b. Write up furnished by the Agency on the methodology to be followed for the tasks described at 1-6 above.
    c. Sample Report of similar work undertaken in the past.
    d. Earnest Money Deposit (EMD) for a sum of Rs. 10,000/- (Rupees Ten Thousand Only) in the form of Demand Draft in the name of the **Secretary, NCERT payable at New Delhi** or Bank Guarantee or **Fixed Deposit Receipt**.
3. The **Financial Quotation** must be submitted in the proforma attached at **Annexure 'E'**. This must be placed in a separate envelope super-scribed "**Financial Quotation**".
4. Both the envelopes containing the Technical and Financial Quotations must be placed in a third envelope addressed to the **Head, Educational Survey Division (ESD), NCERT, Room No. 7, 4th floor, Zakir Hussain Block, Sri Aurobindo Marg, New Delhi 110 016.**
5. The proforma for the Financial Quotation must be completed without any alterations to its format and no substitutes shall be accepted. All blank spaces shall be filled in with the information requested.
6. **The last date for receipt of Quotation is June 14, 2013 up to 03.00 PM.** Any bid received after the deadline for submission of bid prescribed by the NCERT will be rejected and/or returned unopened to the bidder.
7. All quotations submitted without EMD will be rejected. Firms exempted from EMD as per government order may enclose the copy of the government order in support of the exemption. The EMD of unsuccessful firms will be refunded within two weeks of finalizing the quotation. The EMD of the successful firm will be discharged when the contract is signed and performance security is paid.
8. There will be a **Pre-bid Meeting on June 10, 2013 at 11.00 AM** in the Room No. 20, 4th Floor, Educational Survey Division, Zakir Hussain Block, NCERT, Sri Aurobindo Marg, New Delhi 110016 in order to provide the bidders an opportunity to seek clarifications on all aspects of the Quotation Documents. Detailed proceedings of the clarifications sought and given during the Pre-bid meeting will be drawn and circulated.

9. The **Technical Quotations will be opened on June 17, 2013 at 3.00 PM** in the Room No. 20, 4th Floor, Educational Survey Division, Zakir Hussain Block, NCERT, Sri Aurobindo Marg, New Delhi 110016 in the presence of those interested firms who choose to attend the opening of technical quotations.
10. The **Financial Quotation will be opened on June 18, 2013 at 11:00 AM**.
11. The quotation shall remain valid for 30 days after the date of bid opening.
12. The quotation should be clear and without any conditions. Conditional quotations will be rejected.
13. Any delay, even postal delay, in receipt of the quotation would be considered late submission of quotation and rejected. The Quotation must be addressed/handed over to the addressee at Sl. No. 4 above. Mere handing over of the Quotation at the Reception Counter or at any other counter or room or person shall not be considered submission of Quotation.
14. The firm submitting the quotation shall bear all costs associated with the preparation and submission of his/her Quotation, and NCERT, will in no case be responsible or liable for these costs, regardless of the conduct or outcome of the bidding process.
15. At any time prior to the deadline for submission of Quotation, NCERT may amend the terms and conditions by issuing an addendum. The amendment will be uploaded on NCERT website www.ncert.nic.in. The amendment will be binding on all the Firms. In order to afford Prospective Bidders reasonable time in which to take the amendment into account in preparing their Bid, the Purchaser may, at its discretion, extend the deadline for the submission of Bids.
16. Educational Survey Division (ESD), NCERT will notify the award of the contract to the successful firm.
17. Within 4 (four) days of the receipt of notification of award from the NCERT, the successful firm shall furnish the Performance Security i.e. 7% of the contract amount in the form of Bank Guarantee or any short term deposit endorsed in the name of NCERT, Earnest Money shall be forfeited if the successful bidder fails to sign the formal agreement within 7 days from the date of intimation to that effect.
18. Failure of the successful Bidder to comply with the requirement shall constitute sufficient grounds for annulment of the award and forfeiture of the earnest money, in which event NCERT may make the award to the next lowest evaluated bidder or call for new quotations.
19. NCERT reserves the right to reject any quotation without assigning any reason.
20. The firm which is awarded the contract will take delivery of the raw data from ESD, NCERT, New Delhi within 24 hours of signing the contract. NCERT will not bear the expenses involved in delivery of raw data. The firm awarded the contract will submit signed receipts for raw data received.
21. The firm awarded the contract shall strictly adhere to the time schedule stipulated in Table-4. Any delay in the completion of tasks as stipulated in Table-4 shall entail delay liability equal to 0.5% of the contract amount per day of delay. However, if there is delay on the part of the ESD in supplying the raw data to the firm or in completing the scoring and range checks, such delay liability will not be applicable for the period of the delay caused by the ESD.
22. The firm awarded the contract will be required to keep the supplied material safe and in proper order till the analysis is over. After completion of work, the firm shall return the raw data to NCERT within one month at its own expense.

Head, Educational Survey Division
4th Floor, Zakir Hussain Block
National Council of Educational Research and Training (NCERT)
Sri Aurobindo Marg, New Delhi 110 016.

Source: National Council of Educational Research and Training, (2012). Quotation for Analysis of Data of National Achievement Survey.
Accessed at www.ncert.nic.in/announcements/tendors/pdf_files/NAS_C3.pdf

APPENDIX H

National Press Release for Trial Urban District Assessment 2013

NEWS RELEASE
Embargoed: Hold for release until December 18, 2013, at 1 p.m. EST
CONTACT: Stephaan Harris, (202) 357-7504, stephaan.harris@ed.gov

A decade of progress for urban districts in mathematics and reading
But most TUDA scores for fourth- and eighth-grade students lower than the nation

WASHINGTON—Ten years after The Nation's Report Card began measuring progress in America's urban school districts, the 2013 Trial Urban District Assessment (TUDA) shows that most districts that participated in the first reading or math assessments scored higher this year at both grades 4 and 8, and none of the participating districts scored lower than in the first testing year. The District of Columbia Public Schools was the only one of the 21 districts that participated this year to show gains in both mathematics and reading at both grades compared with 2011. In Los Angeles, scores improved in reading at both grades, and in mathematics at grade 4.

The Nation's Report Card: 2013 Mathematics and Reading—Trial Urban District Assessment, reports the achievement of public school students in 21 urban districts on the National Assessment of Educational Progress (NAEP). In addition to providing national and state-level results for context, the urban district assessment includes findings for the nation's large cities (based on the combined scores of all cities in the nation with populations of 250,000 or more—including the participating districts). Some 30 percent of America's students—about 15 million in all—attend schools in urban districts, including those not participating in TUDA. There is a great deal of racial and ethnic diversity throughout the districts, where most of these students are eligible for free and reduced-price lunches.

"Anyone interested in the state of our nation's education should start by looking at progress in these urban districts, which face a concentration of the challenges all schools grapple with to some degree," said David P. Driscoll, chair of the National Assessment Governing Board, which sets policy for NAEP. "By volunteering to be part of TUDA, these districts gain insights and data they can use to focus their academic efforts."

Average reading and mathematics scores for fourth- and eighth-grade students in most TUDA districts, even those that have improved, were lower than the average scores for students in their home states and the nation. When participating urban districts are compared with large cities nationally, more districts score lower than their city peers in reading than in math. For example, in grade 4 math, nine districts scored lower than the average for large cities. In grade 4 reading, 12 districts scored lower than large cities nationally.

The 2013 TUDA results are based on representative samples of 1,100 to 2,300 public school students at grade 4 and 900 to 2,100 public school students at grade 8 in each participating urban district. NAEP attempts to include in its assessments a highly representative sampling of

students, and counts as a factor the percentage of participating students who have disabilities or are English language learners. The District of Columbia Public Schools, for example, at both grades in math, and at fourth grade in reading, included more than 85 percent of its students with disabilities and those learning English—a percentage that exceeds the standard set by the Governing Board as being representative.

Notable progress in closing gaps in achievement in urban districts includes:

- Black, Hispanic and white students in Los Angeles scored higher in 2013 than in 2011 in mathematics at grade 4.
- Black, Hispanic and white students in the District of Columbia scored higher in 2013 than in 2011 in reading at grade 8.
- Students who are eligible for free or reduced-price lunches increased their average scores from 2011 to 2013 in at least one subject and grade combination in eight districts (Atlanta; Baltimore City; Charlotte, N.C.; Chicago; Dallas; D.C.; Fresno, Calif.; and Los Angeles).

"Every district has its own story, but as a whole over the last 10 years all of the districts are improving," Driscoll said. "In general, though, these scores are too low, and that should concern everyone. TUDA matters because these school systems need our attention more than ever before."

Scores fall on a 0-500 scale, and are divided into achievement levels described as *Basic* (partial mastery of the knowledge and skills needed at that grade)*, Proficient* (solid academic performance) and *Advanced* (superior work)*.* The National Center for Education Statistics, in partnership with the Governing Board and Council of the Great City Schools, created TUDA in 2002 to support the improvement of student achievement in the nation's large urban districts. The TUDA measures educational progress within participating large urban districts. Reading results were first reported for six urban districts in 2002, and mathematics results were first reported in 2003 for 10 districts. Since 2002, urban districts have been added, culminating in the 21 districts that participated in both 2011 and 2013.

This report card is the second to be published in an interactive online report that allows searches using multiple variables within districts and for comparison against other regions. The "district profiles" pages include, for example, the performance gaps by race/ethnicity, gender, and eligibility status for the National School Lunch Program. It also includes classroom context, such as how much time teachers spend teaching a subject compared with the corresponding information at the state and national levels. The report website also features a video to help people understand the multiple ways the new site allows searches for hundreds of findings.

###

The National Assessment of Educational Progress (NAEP) is a congressionally authorized project sponsored by the U.S. Department of Education. The National Center for Education Statistics, within the Institute of Education Sciences, administers NAEP. The Commissioner of Education Statistics is responsible by law for carrying out the NAEP project.

The National Assessment Governing Board is an independent, bipartisan board whose members include governors, state legislators, local and state school officials, educators, business representatives and members of the general public. Congress created the 26-member Governing Board in 1988 to set policy for NAEP.

APPENDIX I

Use of Assessment Results

Selected Countries that Used National Assessment Results in Reviewing the Education System

Country	Examples of some claimed uses
Argentina	Instituted a program of school inspection
Bolivia	Linked assessment data to a program for child nutrition
Burkina Faso	Provided input for country analysis
Cuba	Strengthened preschool and early childhood care programs
Kenya	Led to benchmarks for providing facilities
Kuwait	Provided support for the policy of introducing classroom libraries
Malawi	Provided input for reform program
Mauritius	Used data to support national sector study
Namibia	Used by national commission
Nepal	Supported major government reform program
Niger	Provided input for country analysis
Sri Lanka	Provided input for national sector strategy for education
Uganda	Used to prepare educational reform program
Uruguay	Used to support a policy of expanding an equity program for full-time schools
Vietnam	Used to establish benchmarks for providing facilities (desks per pupil, books per pupil)
Zanzibar (Tanzania)	Used in review of educational policies, standards, and benchmarks
Zimbabwe	Used in commission review

Source: Table reproduced from Kellaghan and Greaney (2009)

APPENDIX J

Advantages and Disadvantages of using Census Based Large Scale Assessments

Advantages	Disadvantages
Focuses on what are considered important aspects of education.	Tends to lead to neglect of subject areas that are not tested.
Highlights important aspects of individual subjects.	Tends to lead to neglect of aspects of subjects that are not tested (such as oral fluency in language).
Helps ensure that students reach an acceptable standard before promotion.	Has contributed to early dropout and non-promotion.
Allows for direct comparisons of schools.	Leads to unfair ranking of schools where different social backgrounds are served and where results are not significantly different.
Builds public confidence in the performance of the system.	Has led to cheating during test administration and to subsequent doctoring of results.
Puts pressure on students to learn.	Tends to emphasize memorization and rote learning.
Results in some schools and students raising test performance levels.	Improved performance may be limited to a particular test and will not be evident on other tests of the same subject area.
Allow parents to judge the effectiveness of individual schools and teachers.	Leads to unfair assessment of effective- ness on the basis of test score performance rather than taking into account other established factors related to learning achievement.
Tends to be popular with politicians and media.	Seldom holds politicians accountable for failure to support delivery of educational resources.

Source: Table reproduced from Kellaghan and Greaney (2008)

APPENDIX K

Sample Supporting Questionnaire

Structure of Supporting Background Questionnaire

Questionnaire	Variable
Student	▪ Gender, age, and language background (all usually collected on the front of the test booklet) ▪ Educational background, such as years at school and periods away from school ▪ Opportunities to attend school ▪ Expectations of success and personal or family attitudes about the value of school ▪ Perceptions of classroom environments, such as sense of safety, friendliness of other students, or support from teachers
Parent	▪ Nationality, gender, and language background ▪ Home environment, such as access to books, desks, and lights ▪ Family background, such as education of parents and language spoken at home ▪ Attitudes toward education, such as commitment to sending children to school, perceptions of the value and relevance of education, or perceptions of the quality of education ▪ Attention to homework and study resources provided at home for children ▪ Affordability and accessibility of education for children ▪ Expectations of educational achievement for children ▪ Involvement with schools, such as participation in the classroom or on committees ▪ Nature of school reports about children's progress and their value ▪ Financial support for school in the form of payment for textbooks and fees

Questionnaire	Variable
Teacher	▪ Gender and age ▪ First language ▪ Teaching conditions, such as class size, access to resources, percentage of students who have textbooks, access to replacement teachers when sick, and assistance with challenging students ▪ Educational experience, teacher qualifications, and number of year in this school ▪ Professional engagement with learning, such as access to and interest in professional development, interest in teaching, and time spent preparing for classes ▪ Availability of instructional support through classroom visits by head-teachers, school inspectors, or supervisors ▪ Teaching methodology, such as language of instruction, use of assessment, and style of teaching ▪ Satisfaction with working conditions, such as tenure, rates of pay, and level of supervision ▪ Relationship with the school community, such as interactions with parents, involvement in school committees, and participation in local community events ▪ Distance from teacher's home to school
Head Teacher	▪ Gender and age ▪ Educational and management experience and qualifications ▪ School environment, such as quality of buildings and facilities, as well as availability of resources ▪ School records, such as fluctuations in student numbers, the extent of student or teacher absenteeism, and the frequency of students changing schools ▪ Professional engagement with school leadership, such as access to and interest in professional development and interest in education ▪ Leadership style and use of time ▪ Satisfaction with working conditions, such as tenure, rates of pay and level and frequency of supervision ▪ Relationship with school community, such as interactions with parents and participation in local community events.

Source: Table reproduced from Anderson and Morgan (2008)

APPENDIX L

Case Study 3: Struture of Pratibha Parv Workday

शिक्षा का अधिकार
सर्व शिक्षा अभियान
सब पढ़ें सब बढ़ें

राज्य शिक्षा केन्द्र

पुस्तक भवन, बी-विंग, अरेरा हिल्स, भोपाल—462 011
दूरभाष : (0755) 2768390, 91, 92, 94, 95 फैक्स : 2552363, 2760561

/राशिके/मूल्यांकन/2013/10723 भोपाल, दिनांक 5.12.13

समय–सीमा

ति,

कलेक्टर,
समस्त जिले, म.प्र.

विषय– "प्रतिभा पर्व" 2013–14 के अंतर्गत मूल्यांकन।

संदर्भ - राशिके के पत्र क्र.10207 दि.19.10.12, क्र.1086 दि .07.02.2 13 एवं 2307 दि. 01.03.13.

प्रतिभा पर्व 2013–14 के अंतर्गत प्रदेश के समस्त शासकीय प्राथमिक एवं माध्यमिक शालाओं में आगामी दिनांक 10 जनवरी, 2014 को एक साथ प्रतिभा पर्व का आयोजन किया जा रहा है। प्रतिभा पर्व के अंतर्गत शालेय व्यवस्था संचालन की स्थिति व बच्चों की शैक्षिक उपलब्धियों का विभागीय स्तर पर जायजा लिया जाएगा, ताकि सुधार हेतु प्रयास किए जा सकें। इसमें शाला प्रबंधन समिति के सदस्यों, पंच–सरपंच–पालक/जनप्रतिनिधि आदि की सहभागिता भी होगी।

दिनांक 13 दिसम्बर, 2013 को राज्य शिक्षा केन्द्र स्तर पर प्राचार्य डाइट/डीआरसी व एसीसी अकादमिक का प्रशिक्षण रखा गया है, जिसमें उन्हें प्रश्नपत्रों व प्रपत्रों की सीडी प्रदाय की जाएगी। सीडी प्राप्त होने के पूर्व मुद्रण संबंधी समस्त औपचारिकताएँ यथा–कोटेशन आमंत्रण, न्यूनतम दर के कोटेशन को मुद्रण हेतु आदेश जारी करना आदि पूरी कराने का अनुरोध है, ताकि मुद्रण कार्य समय सीमा में पूर्ण किया जा सके। गतवर्ष के निर्देशों में आंशिक संशोधन किया गया है, शेष निर्देश प्रतिभा पर्व वर्ष 2012 13 के संदर्भित पत्रानुसार रहेंगे।

संशोधित निर्देश वर्ष 2013–14 निम्नानुसार रहेंगे–

1. मुद्रण दर

आकार	जी.एस.एम.	पृष्ठ संख्या	अधिकतम व्यय सीमा	प्रश्नपत्र–सह–उत्तरपुस्तिका/प्रपत्र	कलर
ए–4 साइज	60	1 पेज	40 पैसे	डाटा एंट्री प्रपत्र	ब्लेक
ए–3 साइज	60	1 पेज	70 पैसे	शैक्षिक मूल्यांकन प्रपत्र–1	ब्लेक
ए–4 साइज	60	2 पेज आगे–पीछे	60 पैसे	कक्षा–1, 2, 3, 4	ब्लेक
ए–4 साइज	60	4 पेज आगे–पीछे	95 पैसे	कक्षा–5	ब्लेक
ए–4 साइज	60	6 पेज आगे- पीछे	130 पैसे	कक्षा–6, 7, 8	ब्लेक

नोट–कृपया उक्त कार्रवाई में म.प्र. भंडार क्रय नियमों का पूर्णतः पालन करते हुए समय–सीमा में कार्रवाई सुनिश्चित की जाए।

2. कक्षा 1–4 तक के प्रश्न पत्रों का स्वरूप यथावत् रहेगा, परन्तु कक्षा 5 से 8 तक के प्रश्नपत्रों के स्वरूप में आंशिक परिवर्तन किया गया है। कक्षा–5 से 8 तक के प्रश्नपत्रों में प्र.1–5 बहुविकल्पीय, प्र.6–7 रिक्त स्थान की पूर्ति/दिए गए शब्दों का वाक्य में प्रयोग व प्र.8–10 तक एक शब्द/एक वाक्य में उत्तर स्वरूप के तथा अनुच्छेद पर आधारित रहेंगे, लेकिन गणित के सभी प्रश्न बहुविकल्पीय स्वरूप के रहेंगे। प्रश्नपत्र नवम्बर माह तक के पाठ्यक्रम पर आधारित होंगे। तदनुसार बच्चों की तैयारी करवाने हेतु जिले की शासकीय प्राथमिक व माध्यमिक शालाओं को अनिवार्यतः निर्देशित किया जाए।
3. सत्यापन कार्य हेतु अधिकारियों की ड्यूटी इस तरह लगाई जाए कि वे अपने कार्यस्थल से 8 कि.मी. की परिधि वाली शालाओं में ही सत्यापन कार्य संपादित कर सकें, ताकि टीए–डीए की राशि में मितव्ययिता की जा सके। 8 कि.मी. से अधिक की दूरी की शालाओं का सत्यापन कार्य करने वाले अधिकारियों को अपने संबंधित कार्यालय/विभाग से ही टीए–डीए आहरित करने हेतु निर्देशित किया जाए।
4. प्रतिभा पर्व का व्यापक प्रचार–प्रसार किया जाए। इस कार्य में शाला प्रबंधन समिति व शिक्षकों के साथ ग्राम पंचायत/नगरीय निकाय का सहयोग लिया जाए। इसके लिए शाला प्रबंधन समिति सदस्यों, सरपंच/पंच/पार्षद को पत्र लिखकर यह अनुरोध किया जाए कि वे प्रतिभा पर्व से एक सप्ताह पहले ग्राम/वार्ड में चौपाल लगाकर व डोंडी पिटवाकर माता–पिता/अभिभावकों से समस्त छात्र–छात्राओं को प्रतिभा पर्व के दिन शाला में उपस्थित रहने को कहा जाए तथा सरपंच–पंच/जनप्रतिनिधि भी प्रतिभा पर्व के दिन शाला का अवलोकन करें।
5. प्रतिभा पर्व मूल्यांकन दिवस को स्पेशल मध्याह्न भोजन (मिड–डे–मील) उपलब्ध कराया जाए।
6. प्रतिभा पर्व आयोजन दिनांक को शाला में की जाने वाली गतिविधियों का क्रम चार्ट (परिशिष्ट–1) व प्रतिभा पर्व का गतिविधि केलेण्डर (परिशिष्ट-2) संलग्न है, जिसके अनुसार जिले में कार्यवाई अपेक्षित है।

संलग्न :- उपर्युक्तानुसार।

रश्मि अरुण शमी
आयुक्त, राज्य शिक्षा केन्द्र एवं
पदेन सचिव, म.प्र. शासन, स्कूल शिक्षा।

पृ.क्र./राशिके/मूल्यांकन/2013/10724 भोपाल, दिनांक 5.12.13

प्रतिलिपि–

1. निज सचिव, माननीय मंत्रीजी, म.प्र.शासन, स्कूल शिक्षा (म.प्र.) की ओर सूचनार्थ।
2. निज सचिव, माननीय राज्य मंत्रीजी, म.प्र.शासन, स्कूल शिक्षा (म.प्र.) की ओर सूचनार्थ।
3. प्रमुख सचिव, म.प्र. शासन, स्कूल शिक्षा (म.प्र.) की ओर सूचनार्थ।
4. प्रमुख सचिव, म.प्र. शासन, आदिवासी विकास (म.प्र.) की ओर सूचनार्थ।
5. आयुक्त, लोक शिक्षण भोपाल (म.प्र.) की ओर आवश्यक कार्रवाई हेतु
6. आयुक्त, आदिवासी विकास विभाग, भोपाल (म.प्र.) की ओर आवश्यक कार्रवाई हेतु।
7. आयुक्त, पंचायत एवं ग्रामीण विकास विभाग, भोपाल (म.प्र.) की ओर आवश्यक कार्रवाई हेतु।
8. सीईओ, जिला पंचायत,समस्त जिले, म.प्र. की ओर आवश्यक कार्रवाई हेतु
9. संयुक्त संचालक, लोक शिक्षण, समस्त संभाग (म.प्र.) की ओर आवश्यक कार्रवाई हेतु।
10. प्रभारी, डाइट/डी.आर.सी., समस्त जिले की ओर आवश्यक कार्रवाई हेतु।
11. जिला शिक्षा अधिकारी/राहा,आयुक्त, आदि. विकास विभाग,समस्त जिले की ओर आवश्यक कार्रवाई हेतु।
12. जिला परियोजना समन्वयक,जिला शिक्षा केन्द्र, समस्त जिले,म.प्र. की ओर आवश्यक कार्रवाई हेतु।
13. म.प्र एजुकेशन पोर्टल पर अपलोड करने हेतु।

आयुक्त, राज्य शिक्षा केन्द्र एवं
पदेन सचिव, म.प्र. शासन, स्कूल शिक्षा

परिशिष्ट 1

प्रतिभा पर्व वर्ष 2013–14

प्रतिभा पर्व आयोजन दिनांक को शाला में की जाने वाली गतिविधियों का क्रमचार्ट

क्र.	गतिविधि	निर्धारित समय
1.	–शालेय व्यवस्था प्रपत्र भरना एवं शैक्षिक मूल्यांकन प्रपत्र–1 व 2 में आवश्यक प्रविष्टियां जैसे-- बच्चों के नाम, स्कूल का नाम, डाइस कोड आदि पूर्ण करना, –मूल्यांकन संबंधी आवश्यक तैयारी करना।	प्रतिभा पर्व आयोजन दिनांक के पूर्व
2.	–समस्त स्कूल स्टाफ का शाला में प्रातः 10बजे उपस्थित होना। –बच्चों की उपयुक्त बैठक व्यवस्था कराना,	प्रातः 10 बजे से 10:30 बजे
3.	–मूल्यांकन संबंधी समस्त औपचारिकताएं पूरी करना। –प्रश्नपत्र के सील बंद लिफाफे खोलना,	प्रातः 10:30 बजे से 10:45 बजे
4.	विद्यार्थियों के प्रश्नपत्र वितरण व्यवस्था करना।	प्रातः 10:45 से 11 बजे
5.	बच्चों का लिखित/मौखिक मूल्यांकन।	प्रातः 11:00 से दोपहर 1:00 बजे
6.	मध्याह्न भोजन (साबुन से हाथ धुलाना, पंक्तिबद्ध बिठाकर भोजन कराना)	दोपहर 1:00 से दोपहर 2:00 बजे
7.	प्रधानाध्यापक, एसएमसी सदस्य एवं जनप्रतिनिधियों की उपस्थिति में **बाल केबिनेट** (बालकेबिनेट का गठन व क्रियान्वयन निर्देश राशिके पत्र क्र.8923 दि.17.9.2012) द्वारा बच्चों की बालसभा (सांस्कृतिक/ साहित्यिक/ सृजनात्मक/ खेलकूद/ वैज्ञानिक गतिविधियां) का आयोजन कराना।	दोपहर 2:00 से अपराह्न 4:30 बजे
8.	शिक्षकों द्वारा उत्तरपुस्तिकाओं का मूल्यांकन, मूल्यांकन संबंधी प्रपत्र भरना।	दोपहर 2:00 से अपराह्न 4:30 बजे
9.	परिणाम की घोषणा एवं डाटा एंट्री प्रपत्र शाला में आए सत्यापनकर्ता अधिकारी को सौंपना।	अपराह्न 4:30 बजे

परिशिष्ट–2

प्रतिभा पर्व वर्ष 2013–14 का गतिविधि कलेण्डर

क्र	क्या	प्रक्रिया	कब
1	प्रतिभा पर्व के निर्देश जारी करना	राज्य शिक्षा केन्द्र	5 दिसम्बर, 2013
2.	प्रतिभा पर्व राज्य स्तरीय प्रशिक्षण व मुद्रण हेतु सीडी प्रदाय की डाइट/डीआरसी प्राचार्य व एपीसी अकादमिक को सूचना देना	डीपीसी द्वारा सूचना	10 दिसम्बर, 2013
3.	राज्य स्तरीय प्रशिक्षण	राज्य शिक्षा केन्द्र	13 दिसम्बर, 2013
4	जिला कलेक्टर/सीईओ के हस्ताक्षर से सत्यापनकर्त्ता अधिकारियों के नियुक्ति आदेश जारी करना (पत्र के बिन्दु क्र.3 के अनुसार)	10 जनवरी को 1 शाला (प्राथमिक+माध्यमिक एक भवन में संचालित होने पर दोनों शालाएं) 11 से 15 जनवरी के बीच 3–4 शालाएं	16 दिसम्बर, 2013
5	जिला स्तरीय एक दिवसीय प्रशिक्षण (सीटीई, मनोविज्ञान मह., PPTI, SISE व डाइट स्टाफ, संकुल प्राचार्य, एपीसी, बीआरसी को)	डाइट/डीआरसी प्राचार्य व एपीसी अकादमिक द्वारा	18 से 18 दिसम्बर, 2013
6.	संकुल/विकासखंड स्तरीय एक दिवसीय प्रशिक्षण (बीईओ, बीएसी, बीजीरी, जनशिक्षक, प्रधानाध्यापक को)	डाइट/डीआरसी प्राचार्य, स्टाफ व एपीसी अकादमिक द्वारा	19–21 दिसम्बर, 2013
7.	प्रश्नपत्रों को मुद्रित कराना (प्रश्नपत्रों को शालावार, कक्षावार सुरक्षित तरीके से मुद्रित कराना व लिफाफे मे सील बद करना)	डीपीसी	25 दिसम्बर, 2013 तक
8.	प्रश्नपत्रों के सीलबंद लिफाफे तथा शैक्षिक मूल्यांकन प्रपत्र–1 व डाटा एंट्री प्रपत्र बीआरसी को सौंपने हेतु तैयारी करना	डीपीसी	3 जनवरी, 2014 तक
9.	प्रश्नपत्रों के शालावार सीलबंद लिफाफे तथा शैक्षिक मूल्यांकन प्रपत्र–1 व डाटा एंट्री प्रपत्र प्राप्त करना	बीआरसी	5 जनवरी, 2014 तक
10.	प्रश्नपत्रों के शालावार सीलबंद लिफाफे तथा शैक्षिक मूल्यांकन प्रपत्र–1 व डाटा एंट्री प्रपत्र जनशिक्षकों को सौंपना	बीआरसी	6 जनवरी, 2014 तक
11.	प्रश्नपत्रों के शालावार सीलबंद लिफाफे तथा शैक्षिक मूल्यांकन प्रपत्र–1 व डाटा एंट्री प्रपत्र प्रधानाध्यापक को सौंपकर पावती लेना	जनशिक्षक	7 जनवरी, 2014 तक
12.	प्रशिक्षण में अनुपस्थित प्रधानाध्यापक को प्रशिक्षण देना (संबंधित शालाओं में जाकर)	जनशिक्षक	8 जनवरी, 2014 तक
13.	शैक्षिक मूल्यांकन प्रपत्र–1 व डाटा एंट्री प्रपत्र का सत्यापन करना	सत्यापनकर्त्ता अधिकारी	10–15 जनवरी, 2014 तक
14.	सत्यापित डाटा एंट्री प्रपत्र प्राप्त करना	बीआरसी	11 से 15 जनवरी, 2014 के मध्य
15.	डाटा एंट्री प्रपत्र की जानकारी पोर्टल पर निर्धारित मॉड्यूल में डाटा एंट्री कराना	बीआरसी	25 जनवरी, 2014 तक

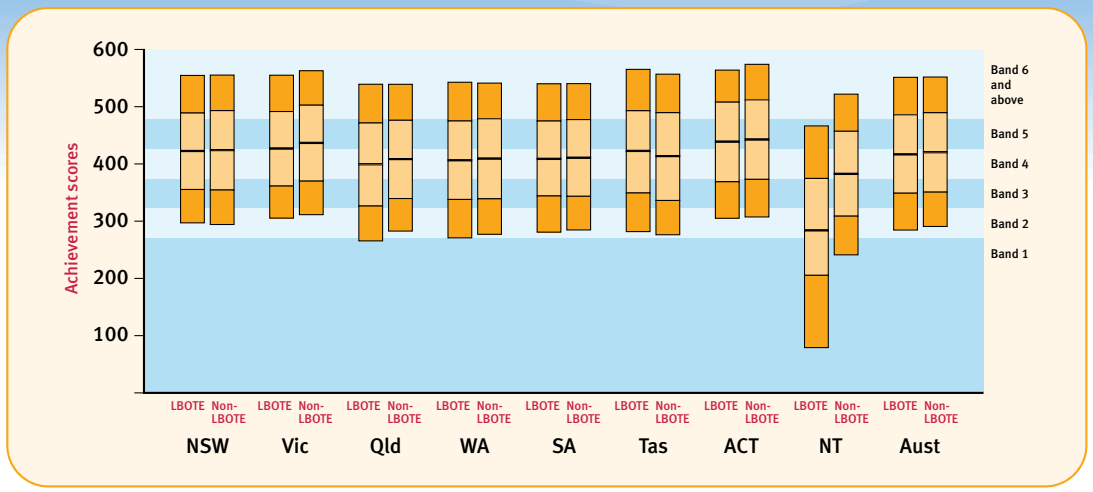
Source: Government of Madhya Pradesh, Pratibha Parv Circular – 2013/10723 accessed at: http://www.educationportal.mp.gov.in/Portal/Public/View.ashx?id=16113&Mode=Circular

APPENDIX M

Case Study 6: Sample from the 2013 NAPLAN Report

NAPLAN Year 3 Reading

Figure 3.R4: Achievement of Year 3 Students in Reading, by LBOTE Status, by State and Territory, 2013.



	NSW	Vic	Qld	WA	SA	Tas	ACT	NT	Aust
LBOTE Mean scale score / (S.D.)	422.6 (78.1)	426.8 (76.1)	399.4 (83.5)	406.4 (82.1)	408.8 (81.3)	422.7 (84.3)	438.8 (80.2)	283.9 (111.7)	416.7 (82.4)
Non-LBOTE Mean scale score / (S.D.)	424.2 (79.6)	436.7 (76.6)	408.5 (78.3)	409.2 (80.7)	410.7 (78.2)	413.7 (86.2)	442.6 (81.3)	382.6 (86.5)	420.5 (79.7)

Table 3.R4: Achievement of Year 3 Students in Reading, by LBOTE Status, by State and Territory, 2013.

State/ Territory	LBOTE status	Below national minimum standard (%)		At national minimum standard (%)	Above national minimum standard (%)				At or above national minimum standard (%)
		Exempt	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6 and above	
NSW	*LBOTE*	2.2	1.9	7.7	17.5	24.3	22.9	23.5	95.9
	Non-LBOTE	1.4	2.2	8.2	16.7	23.0	23.2	25.4	96.5
Vic	*LBOTE*	3.6	1.3	6.6	16.5	24.3	23.6	24.1	95.1
	Non-LBOTE	2.6	1.1	5.6	14.3	22.6	24.6	29.3	96.4
Qld	*LBOTE*	2.8	5.5	12.5	19.9	22.7	19.1	17.4	91.7
	Non-LBOTE	1.4	3.2	10.7	19.9	24.3	21.3	19.2	95.4
WA	*LBOTE*	2.6	4.7	9.9	18.9	24.3	20.9	18.6	92.7
	Non-LBOTE	0.9	4.0	10.4	19.1	23.9	21.6	20.2	95.2
SA	*LBOTE*	4.4	3.6	8.9	18.9	24.1	22.0	18.2	92.0
	Non-LBOTE	1.7	3.2	9.5	19.4	24.5	22.1	19.6	95.1
Tas	*LBOTE*	3.3	3.3	10.1	13.7	20.3	24.7	24.6	93.4
	Non-LBOTE	1.4	3.9	11.1	18.6	21.3	20.2	23.5	94.7
ACT	*LBOTE*	5.6	1.7	5.7	13.2	19.6	23.2	31.1	92.8
	Non-LBOTE	1.4	1.6	5.5	13.0	20.6	24.4	33.5	97.0
NT	*LBOTE*	1.8	42.5	21.6	14.3	9.7	6.0	4.0	55.6
	Non-LBOTE	1.6	9.3	15.0	20.6	22.4	17.6	13.5	89.2
Aust	*LBOTE*	2.8	3.3	8.4	17.5	23.8	22.2	22.1	93.9
	Non-LBOTE	1.7	2.5	8.6	17.4	23.4	22.7	23.8	95.9

Refer to the introduction for explanatory notes and how to read the graph.

APPENDIX N

Case Study 7: SIMCE Dissemination Strategy

SIMCE Dissemination Strategy: Mechanisms, Purposes, Audiences, and Content

Assessment guidelines (since 1988)
Purpose: Provide pedagogical support *Audience*: School principal, pedagogical coordinators and teachers *Content*: (a) Assessment framework and its relationship to the national curriculum (b) Examples of test questions with an analysis of the contents and skills required to answer them correctly *Others*: Distributed to all schools before the assessment (usually in the middle of the school year). Also available online. Publication highly valued by teachers
School report (since 1988)
Purpose: Provide pedagogical support *Audience*: School principal, pedagogical coordinators and teachers *Content*: (a) National-, school-, and class-level mean scores by subject areas and classes tested (b) Differences between school mean scores and mean scores from the previous assessment, from the national mean, and from schools of the same socioeconomic group (c) Percent of students by performance level—advanced, intermediate, beginner (d) Examples of test questions with an analysis of the contents and skills required to answer them correctly (e) Workshop guidelines for the schools to analyse assessment results and set improvement plan *Others*: Distributed to all schools that participated in the assessment once the SIMCE results are released (usually at the beginning of the next school year)
National report (since 2006)
Purpose: Inform policy *Audience*: Decision makers, general public *Content*: (a) National and regional mean scores in subject areas and grades tested (b) Percent of students by performance level—advanced, intermediate, beginner (c) Mean scores by socioeconomic background, gender, public/private school (d) Trends in mean scores across years *Others*: Distributed at the central, regional, and provincial offices of the Ministry of Education. Distributed to persons likely to be interviewed by the media (e.g. university professors)

Newspaper supplement (since 1995)
Purpose: Hold schools accountable *Audience*: Parents, general public *Content*: (a) School mean scores, and mean scores by subject areas and grades tested (b) Differences between school mean scores and mean scores from the previous assessment, from the national mean, and from the mean of schools from the same socioeconomic group *Others*: Published in a newspaper with national and regional coverage. Usually accompanied by rankings of schools
Parent report (since 2005)
Purpose: Hold schools accountable and involve parents in school *Audience*: Parents *Content*: (a) School mean scores, and mean scores by subject areas and grades tested (b) Differences between school mean scores, and between subject area/grade mean scores of schools from the same socioeconomic group (c) Percent of students reaching different performance standards (d) Recommendations to support student learning *Others*: Distributed to parents through the schools once the assessment results are released (usually at the beginning of the school year). Also available online
Online item bank (since 2007)
Purpose: Provide pedagogical support *Audience*: Teachers *Content*: Offers released test questions from all subject areas and target classes. Includes questions from both the national and international assessments *Others*: Teachers can search test questions based on subject area, school cycle, and questions format (multiple choice or open-ended)

Source: Table reproduced from Ramirez (2012). Disseminating and using student assessment information in Chile. Washington, DC: World Bank.

END NOTES

1. Azim Premji Foundation. (2013). Large scale assessments – Guidelines for design, implementation and institutionalization. Prepared for the Ministry of Human Resource Development, New Delhi.
2. Kellaghan, T., & Greaney, V. (2001b). Using assessment to improve the quality of education. Paris: International Institute for Educational Planning.
3. Oza, J., & Bethell, G. (2013). Assessing learning outcomes: policies, progress and challenges. New Delhi.
4. Kellaghan, T., & Greaney, V. (2001b). Using assessment to improve the quality of education. Paris: International Institute for Educational Planning;
5. Kellaghan, T., & Greaney, V. (2004). Assessing student learning in Africa. Washington, DC: World Bank.
6. Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.
7. Alliance for Excellent Education (AEE). (2010). Principles for a comprehensive assessment system. Washington, DC: Author.
8. Eurydice Network. (2009). National testing of pupils in Europe: Objectives, organisation and use of results. Brussels: Education, Audiovisual and Cultural Executive Agency.
9. OECD. (2013). Synergies for better learning. Paris: OECD Publishing.
10. Eurydice Network. (2009). National testing of pupils in Europe: Objectives, organisation and use of results. Brussels: Education, Audiovisual and Cultural Executive Agency.
11. Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.
12. Howie, S. (2004). A national assessment in mathematics within an international comparative assessment: research article. Perspectives in Education, 22 (2).
13. Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.
14. Ibid.
15. Ibid.
16. Ibid.
17. Kirsch, I. (2001). The International Adult Literacy Study (IALS): Understanding what was measured. Princeton, NJ: Educational Testing Service.

18. Mullis, I. V. S., A. M. Kennedy, M. O. Martin, and M. Sainsbury. (2006). PIRLS 2006: Assessment Framework and Specifications. Chestnut Hill, MA: International Study Center, Boston College.

19. Mullis, I. V. S., M. O. Martin, G. J. Ruddock, C. Y. O'Sullivan, A. Arora, and E. Erberber. 2013. TIMSS 2015 Assessment Frameworks. Chestnut Hill, MA: International Study Center, Boston College.

20. Greaney, V., & Kellaghan, T. (2012). Implementing a national assessment of educational achievement. Washington, DC: World Bank.

21. Educational Initiatives. (2010a). Student learning study, status of student learning across 18 states of India in urban and rural schools. Ahmedabad: Educational Initiatives.

22. Ibid.

23. Ibid.

24. Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.

25. Ibid.

26. Anderson P., & Morgan G. (2008). Developing tests and questionnaires for national assessment of educational achievement. Washington, DC: World Bank.

27. Ibid.

28. Ibid.

29. Howie, S & Acana, S. (2012). 'Preparation in Administration in Schools'. In Greaney, V., & Kellaghan, T. (2012). Implementing a national assessment of educational achievement. Washington, DC: World Bank.

30. Educational Initiatives. (2010a). Student learning study, status of student learning across 18 states of India in urban and rural schools. Ahmedabad: Educational Initiatives.

31. Anderson P., & Morgan G. (2008). Developing tests and questionnaires for national assessment of educational achievement. Washington, DC: World Bank.

32. Ibid.

33. Ibid.

34. Azim Premji Foundation. (2013). Large-scale assessments – Guidelines for design, implementation and institutionalization. Prepared for the Ministry of Human Resource Development, New Delhi.

35. Anderson P., & Morgan G. (2008). Developing tests and questionnaires for national assessment of educational achievement. Washington, DC: World Bank.

36. Greaney, V., & Kellaghan, T. (2012). Implementing a national assessment of educational achievement. Washington, DC: World Bank.

37. Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.

38. Greaney, V., Kellaghan, T., & Murray, S.T. (2009). Using the results of national assessment of educational achievement. Washington, DC: World Bank.

39. Pérez, B. A. (2006). Success in implementing education policy dialogue in Peru. Washington, DC: United States Agency for International Development.

40. See http://www. education.ie/servlet/blobservlet/des_cos_preface.htm

41. See http://simce.cl

42. See http://www.erc.ie

43. See http://www.educationportal.mp.gov.in/PratibhaParv/Default.aspx

44. Greaney, V., Kellaghan, T., & Murray, S.T. (2009). Using the results of national assessment of educational achievement. Washington, DC: World Bank.

45. Ilon, L. (1996). Considerations for Costing National Assessments. In National Assessment: Testing the System, ed. P. Murphy, V. Greaney, M. E. Lockheed, and C. Rojas, 69–88. Washington, DC: World Bank.

46. The formats to evaluate each of these indicators is available to view at the Gunotsav website: www.gunotsav.org/gunotsavaheval.aspx

47. Santiago, P., Donaldson, G., Herman, J., & Shewbridge, C. (2011). OECD reviews of evaluation and assessment in education: Australia. OECD.

48. Ibid.

49. Ramirez, M. (2012). Disseminating and using student assessment information in Chile. Washington, DC: World Bank.

REFERENCES

Accountability Initiative, Centre for Policy Research & Assessment Survey Evaluation Research (ASER) Centre (2012). Do Schools Get Their Money? PAISA 2012. New Delhi: Author.

Agencia de Calidad de la Educaicõn, Government of Chile. Agencia de Calidad de la Educaicõn. Retrieved January 24, 2014, from http://www.agenciaeducacion.cl/

Alliance for Excellent Education (AEE). (2010). Principles for a comprehensive assessment system. Washington, DC: Author.

Anderson P., & Morgan G. (2009). Developing tests and questionnaires for national assessment of educational achievement. Washington, DC: World Bank.

Assessment Survey Evaluation Research (ASER) Centre. (2011). Inside primary schools: A study of teaching and learning in rural India. New Delhi: Author.

Assessment Survey Evaluation Research, (ASER) Centre and Pratham. (2005-2013). Annual status of education report. New Delhi: Author.

Australian Council for Educational Research. (2011). PISA 2009 plus results. Victoria, Australia: Author.

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2013). ACARA. Retrieved January 3, 2014, from http://www.acara.edu.au/default.asp

Australian Curriculum, Assessment and Reporting Authority. (2013). NAPLAN national report 2013. Sydney: Author.

Australian Curriculum, Assessment and Reporting Authority. (2011). National Assessment Program. Retrieved January 3, 2014, from http://www.nap.edu.au/

Australian Government, Department of Education. (2011). Smarter schools. Retrieved January 8, 2014 from http://smarterschools.gov.au/.

Azim Premji Foundation. (2013). Large scale assessments – Guidelines for design, implementation and institutionalization. Prepared for the Ministry of Human Resource Development, New Delhi.

Benavot, A., & Tanner, E. (2007). The Growth of National Learning Assessments in the World. Prepared for the Education for All Global Monitoring Report 2008, Education for All by 2015: Will we make it?

Black, P., & William, D. (2001). Inside the black box: Raising standards through classroom assessment. London: NFER Nelson.

Bonamino, A., & Sousa, S.Z. (2012). Three generations of assessments of basic education in Brazil: Interfaces with the curriculum in/of the school. Educação e Pesquisa, 38(2), 373-388.

Braun, H., Kanjee, A., Bettinger, E., & Kremer, M. (2006). Improving education through assessment, innovation and evaluation. Cambridge, MA: American Academy of Arts and Sciences.

Canen, A. (2012). Assessment of schools in Brazil: some reflections. SA-eDUC, 9 (1).

Chandaraiah, S. (November 20, 2013). Personal communication.

Educational Initiatives with Harvard University, Azim Premji Foundation, World Bank and Government of Andhra Pradesh. (2005). Andhra Pradesh randomized evaluation study.

Educational Initiatives. (2010a). Student learning study, status of student learning across 18 states of India in urban and rural schools. Ahmedabad: Educational Initiatives.

Educational Initiatives. (2010b). Quality education study. Ahmedabad: Educational Initiatives.

Eurydice Network. (2009). National testing of pupils in Europe: Objectives, organisation and use of results. Brussels: Education, Audiovisual and Cultural Executive Agency.

Government of India. (2013). Twelfth five year plan (2012-2017) (Vol. 3). New Delhi: Social Sectors, Sage.

Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education. Washington, DC: World Bank.

Greaney, V., & Kellaghan, T. (2009). Implementing a national assessment of educational achievement. Washington, DC: World Bank.

Greaney, V., Kellaghan, T., & Murray, S.T. (2012). Using the results national assessment of educational achievement. Washington, DC: World Bank.

Govinda R. (2007). India country case study. Prepared for the Education for All Global Monitoring Report 2008, Education for All by 2015: Will we make it?

Howie, S. (2004). A national assessment in mathematics within an international comparative assessment: research article. Perspectives in Education, 22 (2).

Ilon, L. (1996). Considerations for Costing National Assessments. In National Assessment: Testing the System, ed. P. Murphy, V. Greaney, M. E. Lockheed, and C. Rojas, 69–88. Washington, DC: World Bank.

International Association for the Evaluation of Educational Achievement (IEA). (2013). Framework and specifications for TIMSS assessment 2015. Chestnut Hill, MA: International Study Center, Boston College.

Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira (INEP), Brazil. INEP portal. Retrieved December 7, 2013 from http://portal.inep.gov.br/

Kellaghan, T., & Greaney,V. (2004). Assessing student learning in Africa. Washington, DC: World Bank.

Kellaghan, T., & Greaney, V. (2001b). Using assessment to improve the quality of education. Paris: International Institute for Educational Planning.

Kirsch, I. (2001). The International Adult Literacy Study (IALS): Understanding what was measured. Princeton, NJ: Educational Testing Service.

Mourshed, M., Chijioke, C., & Barber, M. (2010). How the world's most improved school systems keep getting better. McKinsey & Company.

Mullis, I. V. S., A. M. Kennedy, M. O. Martin, and M. Sainsbury. (2006). PIRLS 2006:

Assessment Framework and Specifications. Chestnut Hill, MA: International Study Center, Boston College.

Mullis, I. V. S., Martin, M.O., Gonzalez, E. J., & Chrostowski, S.J. (2004). TIMSS 2003 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades. Chestnut Hill, MA: International Study Center, Boston College.

Mullis, I. V. S., M. O. Martin, G. J. Ruddock, C. Y. O'Sullivan, A. Arora, and E. Erberber. 2013. TIMSS 2015 Assessment Frameworks. Chestnut Hill, MA: International Study Center, Boston College.

NAS Review & Strategic Planning Committee. (2012). Strategic directions – National assessment system of learning achievement report. New Delhi: Author.

National Council of Educational Research and Training. (2012). National Achievement Survey Class V. New Delhi: Author.

New South Wales Department of Education and Communities. (2011). National Partnership on Literacy and Numeracy. Retrieved January 22, 2014 from http://www.curriculumsupport.education.nsw.gov.au/national/

OECD. (2010). The high cost of low educational performance: The long-run economic impact of improving PISA outcomes. OECD Publishing/PISA.

OECD. (2010). Review of national policies for education: Santa Catarina State, Brazil. OECD Publishing.

OECD. (2011). Strong performers and successful reformers in education: Lessons from PISA for the United States. OECD Publishing.

OECD. (2013). Synergies for better learning. Paris: OECD Publishing.

OECD. (2013). OECD territorial reviews: Brazil, 2013. OECD Publishing.

Oza, J., & Bethell, G. (2013). Assessing learning outcomes: policies, progress and challenges. New Delhi.

Pérez, B. A. (2006). Success in implementing education policy dialogue in Peru. Washington, DC: United States Agency for International Development.

Pritchett, L. (2012, January 5). The first PISA results for India: The end of the beginning. Message posted to http://ajayshahblog.blogspot.in/2012/01/first-pisa-results-for-india-end-of.html

Ramirez, M. (2012). Disseminating and using student assessment information in Chile. Washington, DC: World Bank.

Sankar, V. (2013). Large scale diagnostic assessments – the Gujarat experience. Retrieved from http://www.teachersofindia.org/en/article/large-scale-diagnostic- assessments-gujarat-experience

Santiago, P., Donaldson, G., Herman, J., & Shewbridge, C. (2011). OECD reviews of evaluation and assessment in education: Australia. OECD.

Schulz, W. (February 17, 2014). Personal communication.

Tognolini, J. (January 29, 2014). Personal communication.

United Nations Educational, Scientific and Cultural Organization (UNESCO). (1990). World declaration on education for all: Meeting basic learning needs. Adopted by the World Congress on Education for All. New York: UNESCO.

Photo Credits

Cover: Hemang Mehta, Educational Initiatives

Page 10: 3.2.1 Education Foundation

Page 16: 3.2.1 Education Foundation

Page 28: Akanksha Foundation

Page 36: 3.2.1 Education Foundation

Page 44: 3.2.1 Education Foundation

Page 50: Hemang Mehta, Educational Initiatives

Page 56: Hemang Mehta, Educational Initiatives

Page 64: 3.2.1 Education Foundation

Page 70: Onderwijsgek (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons

Page 73: Kaiksky via morgurfile.com

Page 78: Leo Reynolds via Flickr

ABOUT US

About Central Square Foundation

Central Square Foundation (CSF) is a venture philanthropy fund and policy think tank focused on improving learning outcomes for children from low-income communities, with focus on school education.

We are strictly a philanthropic funding and capacity-building organisation that operates by making early and growth stage grants in education-focused NGOs. In specific we support initiatives around the following themes –

- High quality affordable schools
- Human capital development
- Technology in education
- Accountability

Further details are available on our website - www.centralsquarefoundation.org

About FICCI

Established in 1927, FICCI is the largest and oldest apex business organization in India. Its history is closely interwoven with India's struggle for independence, its industrialization and its emergence as one of the most rapidly growing global economies. A non-government, not-for-profit organization, FICCI is the voice of India's business and industry.

FICCI draws its membership from the corporate sector, both private and public, including SMEs and MNCs; FICCI enjoys an indirect membership of over 2,50,000 companies from various regional chambers of commerce. FICCI provides a platform for sector specific consensus building and networking and as the first port of call for Indian industry and the international business community.

Our Vision

To be the thought leader for industry, its voice for policy change and its guardian for effective implementation.

Our Mission

To carry forward our initiatives in support of rapid, inclusive and sustainable growth that encompasses health, education, livelihood, governance and skill development.

Further details are available on our website - www.ficci.com

CENTRAL SQUARE
FOUNDATION

www.centralsquarefoundation.org
info@centralsquarefoundation.org
Printed by: www.rakeshpress.com

FICCI

www.ficci.com
education@ficci.com