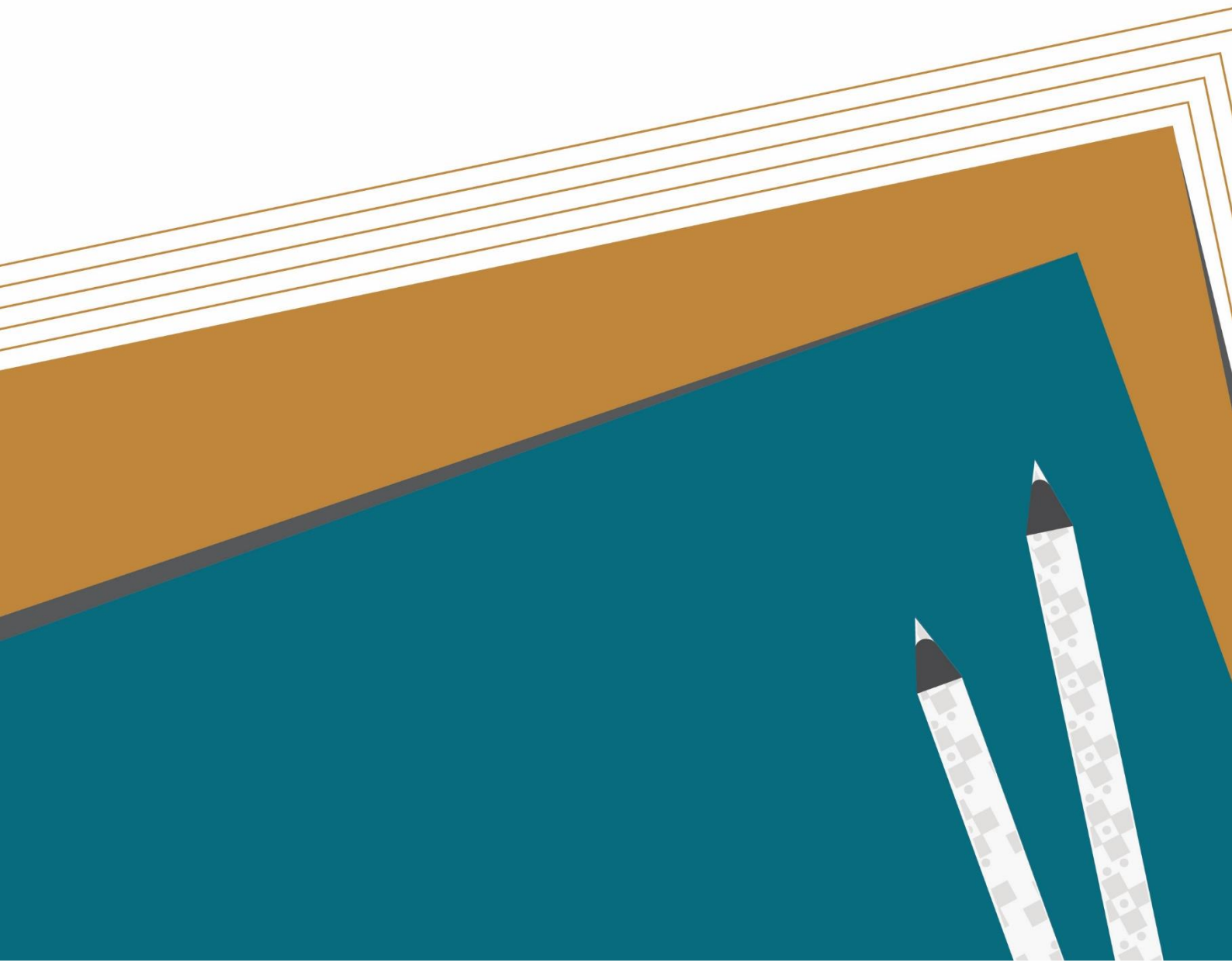


The Strengths and Boundaries of Large-Scale Assessments



The Strengths and Boundaries of Large-Scale Assessmentsⁱ

What are large-scale assessments?

A large-scale assessment (LSA) is a survey designed to measure student achievement in specific curriculum areas, such as reading and mathematics, across an entire education system or within a defined group (e.g., Grade 4 students). The purpose of an LSA is to collect data that helps understand strengths and weaknesses in student learning at a particular point in time, identify trends in performance, and highlight areas for improvement within the education system [1]. LSAs are uniform and standardized in content, administration, timing, and scoring, and are often referred to as standardized tests. While most LSAs are sample-based, an increasing number of countries have adopted a census-based approach in recent decades.

LSAs are considered low-stakes assessments because they have no immediate consequences for individual students, teachers, or schools, and are focused on group-level data. They also maintain anonymity and confidentiality. By gathering background information (e.g., socioeconomic status, language proficiency, and teaching quality), LSAs help policymakers identify achievement gaps, target areas for additional support, and shape policies to address disparities in educational outcomes. LSA results are reported for providing answers to the following questions-

1. Access- What percentage of students have access to school education?
2. Quality- How well are students learning in this education system? Are there particular strengths and weaknesses in students' knowledge and capacities?
3. Equity- Do particular subgroups perform worse than others? Are there disparities, for example, between the performance of boys and girls, students in urban and rural locations, or students from different language groups?
4. Background variables- What factors are associated with student achievement? To what extent does achievement vary with characteristics of learning environment or students home circumstances?
5. Trends- Does student achievement change over time? What factors are linked to changes in student achievement over time?

What has driven the rise of large-scale assessments in India?

The evolution of LSAs in India has been marked by several significant milestones as shown in Figure 1. The National Literacy Mission (NLM), launched by the Ministry of Education (MoE) in 1988, aimed

to reduce illiteracy across the country [2]. The National Literacy Mission (NLM), launched in 1988, aimed to provide functional literacy to 80 million illiterate people aged 15-35, with goals of reaching 30 million by 1990 and an additional 50 million by 1995. The literacy rate in India rose from 52.21% in 1991 to 64.84% in 2001, marking a 12% increase in a decade as indicated by census data. Learning levels were self-reported in census survey- no separate assessment tools were designed and administered to assess literacy rate.

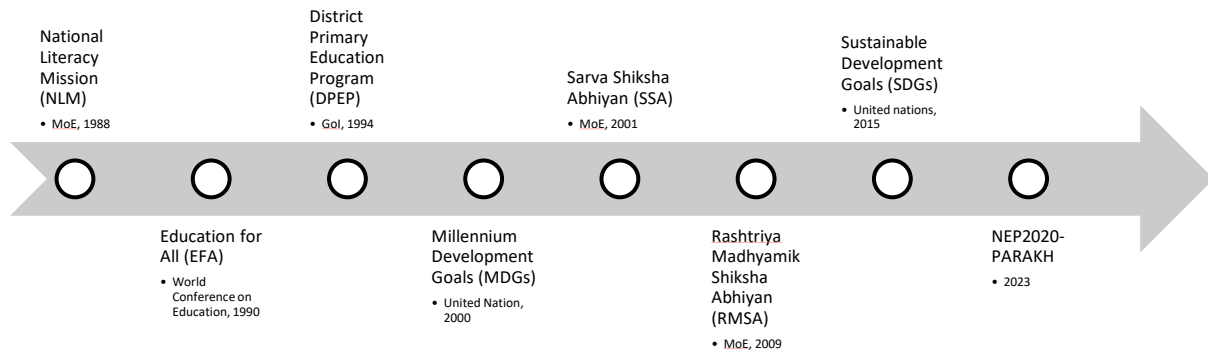


Figure 1: Policies and Initiatives towards Quality Education for All

In 1990, during the International Literacy Year, the World Conference on Education for All in Jomtien, Thailand, gathered 1,500 delegates from 155 countries. India also participated in this conference. They adopted the World Declaration on Education for All and a Framework for Action, calling on all nations to provide universal, adequate basic education. The Declaration emphasized that every individual—child, youth, and adult—should have access to educational opportunities that meet their basic learning needs [3]. As an outcome of this conference, the District Primary Education Programme (DPEP), was launched in 1994, which aimed to revitalize primary education and achieve universal primary education. It adopted an "area-specific approach" with districts as planning units, focusing on local conditions and community participation. DPEP was designed to complement existing central and state primary education schemes by providing additional resources and support. The program also emphasized the role of Village Education Committees (VECs) in ensuring community involvement. The broad objectives of this program were to-

- To provide all children access to primary education through either the formal or non-formal stream.
- To reduce differences in enrollment, dropout rates, and learning achievements among gender and weaker section groups to less than five percent.
- To reduce overall primary dropout rates for all children to less than 10 percent.
- To increase the average achievement rate by 25 percent from the measured baseline level, ensuring basic literacy and numeracy competencies are achieved, along with a minimum of 40 percent achievement in other competencies by all primary education children.

To assess the efficacy of the DPEP program, a baseline achievement survey was conducted in 1994, followed by a midline achievement survey in 1997, and a terminal achievement survey in 2002. These surveys assessed mathematics and language for students in grades 1 and 3 or 4. The results showed that the goal of attaining a minimum of 40% score was achieved in grade 1 in most DPEP

districts. However, it was challenging to determine if the average achievement had increased over the baseline scores, as the competencies assessed in the baseline, midline, and terminal surveys differed due to changes in the curriculum [5].

Following the end of the Cold War, a series of UN-led conferences in the 1990s focused on global issues such as children's rights, nutrition, and women's rights, leading to international commitments for action. In 2000, the Millennium Development Goals (MDGs) were established through the Millennium Declaration to be achieved by 2015. Among the eight goals was Goal 2, which aimed to achieve universal primary education, with Target 2A specifying that by 2015, all children, boys and girls, should complete a full course of primary education [6]. Building on these efforts, the Sarva Shiksha Abhiyan (SSA) was launched by the MoE in 2001 to achieve universal elementary education. The objectives of SSA were that [7]-

- All children complete five years of primary schooling by 2007
- All children complete eight years of elementary schooling by 2010
- Focus on elementary education of satisfactory quality with emphasis on education for life
- Bridge all gender and social category gaps at primary stage by 2007 and at elementary education level by 2010
- Universal retention by 2010

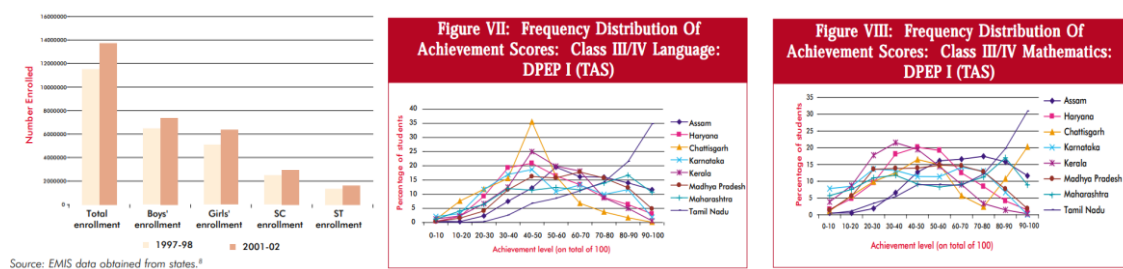


Figure 2: A snapshot of reporting in LSAs conducted during DPEP implementation

This was the period when National Achievement Survey (NAS) [8] and Annual Status of Education Report (ASER) [9] were introduced in our country. NAS is a comprehensive educational assessment survey conducted in India every three years. It is administered by the Department of School Education and Literacy, which operates under the Ministry of Education. Since its inception in 2001, the NAS has been conducted eight times. The primary aim of the NAS is to assess the learning outcomes of students in key subjects like language, mathematics, science, and social science at different grade levels. The test includes the Pupil Questionnaire (PQ), Teacher Questionnaire (TQ) and School Questionnaire (SQ) and Achievement Test (AT). It serves as a diagnostic tool for both the central and state governments to assess the quality of education and design appropriate interventions. It is an important step towards enhancing educational quality and ensuring that every child receives the education they are entitled to under the Right to Education Act.



Figure 3: Cycles of NAS administration

ASER is one of the largest and most well-known citizen-led surveys in India, providing important insights into the quality of education in rural areas. ASER aims to evaluate the basic literacy and numeracy skills of children in rural India, with an emphasis on assessing foundational learning outcomes rather than merely enrolment. The survey also collects data on other aspects of education, such as school facilities, teacher availability, and community participation. The ASER survey has become an important tool for advocacy and decision-making in India's education policy. It is used by government bodies, NGOs, and international organizations to monitor educational progress, identify challenges, and push for necessary reforms to improve education quality in India.



Figure 4: Cycles of ASER administration

The Sustainable Development Goals (SDGs), adopted by the United Nations in 2015, emphasize the need for inclusive and equitable education for all. SDG 4 focuses on the results and outcomes of educational processes. It highlights both the effective acquisition of relevant knowledge, foundational and transferable skills, and competencies as the basis for lifelong learning, as well as the relevance of learning for the world of work and for civic, social, and cultural life. This focus is clearly reflected in the outcome-based global targets and indicators proposed for monitoring progress on SDG 4. Given the centrality of learning outcomes in this new agenda, learning data is increasingly recognized as a critical component of monitoring mechanisms to measure progress towards SDG 4—at the global, regional, and national levels [10].

Most recently, the National Education Policy (NEP 2020) introduced PARAKH (Performance Assessment, Review, and Analysis of Knowledge for Holistic Development), which was established within NCERT to fulfil the core objectives of setting norms, standards, and guidelines, as well as implementing activities related to student assessment. PARAKH has been mandated to organize large-scale achievement surveys to periodically monitor and assess the state of education in the country [11].

Since 2001, nationwide achievement surveys conducted by NCERT have provided valuable insights into India's school education system. Now, as we enter a new era of education with a focus on competency-based learning, PARAKH Rashtriya Sarvekshan 2024 is set to lead the way under the National Assessment Centre—PARAKH. This large-scale national survey will comprehensively evaluate the progress of learners across the Foundational, Preparatory, and Middle Stages, specifically aligned with the objectives outlined in the National Education Policy (NEP) 2020.

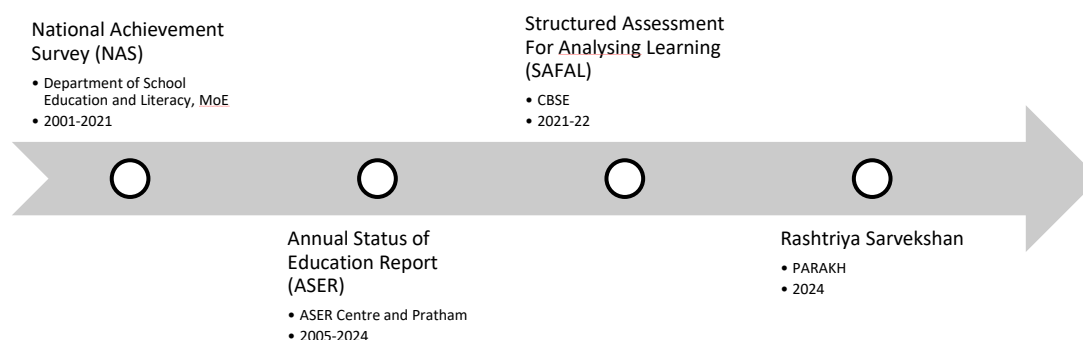


Figure 5: National level LSAs in India

What are the benefits of large-scale assessments?

Understanding learning levels

LSAs can identify disparities in educational achievement across different groups, such as by socioeconomic status, gender, geographic region, or ethnicity. This can help guide targeted interventions to address inequalities. Also, several LSAs focus on assessing core skills, such as literacy, numeracy, and critical thinking, which are essential for further learning and overall development. These insights help identify where students are struggling and the specific skills that need to be reinforced. For example, a study on the loss of learning during the COVID-19 pandemic was conducted by Azim Premji University in January 2021 to examine the extent and nature of the ‘forgetting/regression’ type of learning loss (i.e., knowledge that was previously learned but has now been lost) among children in public schools across primary classes due to school closures [12]. The study focused on assessing four specific abilities in both language and mathematics for students in grades 2 through 6. These abilities were selected because they are foundational for all subsequent learning across subjects. It was concluded that the loss of any one of these skills could have serious consequences for further learning. Also, the extent and nature of learning loss was identified to inform policy and processes as children return to schools.

Informing learning processes

The data from LSAs can provide valuable insights into the effectiveness of different teaching methods, helping educators adapt their strategies to improve student outcomes. If LSAs reveal significant gaps in certain subject areas or competencies, they can prompt curriculum revisions or adjustments to better meet students' needs and ensure they acquire the necessary skills. For instance, the ASER assessment tool highlights the problem of poor foundational learning and aims to promote remedial interventions like Teaching at the Right Level (TaRL). This approach has been rigorously evaluated and proven effective in improving foundational skills in a short time [13].

In 2016, the government of Karnataka adopted and scaled up TaRL, inspired by ASER's findings. Despite high enrolment rates, the annual ASER survey revealed that many Grade 5 students in the

state were unable to read simple texts. For example, in 2016, 42% of Grade 5 students could not read Grade 2-level text in Kannada, the state's official language. In response, the Karnataka Department of State Education Research and Training (DSERT) piloted a remedial program called "Odu [Read] Karnataka" in three districts, aimed at improving literacy and numeracy for Grades 4 and 5 students.

The pilot produced impressive results: the proportion of children reading at Grade 2 level nearly doubled, from 40% to 75%. Encouraged by this success, DSERT, the Sarva Shiksha Abhiyan (SSA), and the Karnataka Department of Public Instruction expanded the program in collaboration with Pratham. Teacher training focused on assessment, grouping students, setting learning goals, and tracking progress, with Pratham providing ongoing support.

Tracking progress on educational goals

Regular LSAs enable tracking of trends in educational performance over time. By comparing results from different years or cycles, it is possible to assess the impact of education reforms, policies, and initiatives. For example, NAS is specifically designed to diagnose the health of the education system. It does so by testing students on few of the learning outcomes prescribed by NCERT for each grade and subject. These learning outcomes are selected from grades 3,5,8 and 10 in key subjects- Language & Mathematics (across all grades), EVS (grades 3 and 5), Science and Social Science (grades 8 and 10). The latest NAS 2021 reported an average learning level of 59% in grade 3, 49% in grade 5, 42% in grade 8 and 36% in grade 10. Compared to NAS 2017, there is a decline in students learning levels across classes 3, 5, 8 & 10 for all the subjects. Average scores in primary stages have dipped considerably as compared to upper primary and secondary stages.

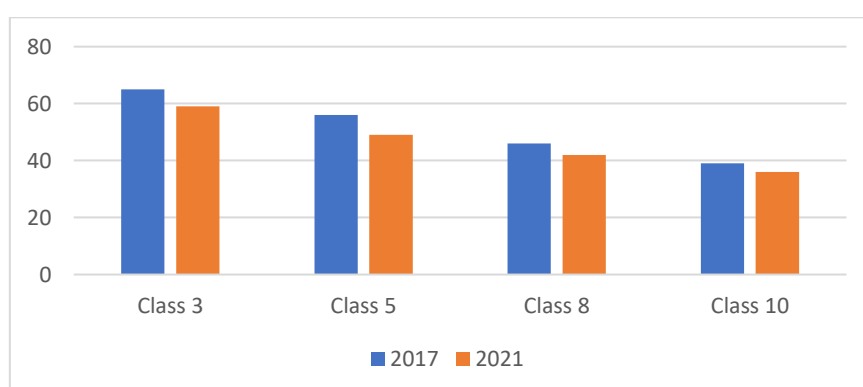


Figure 5: Students performance in NAS 2017 and 2021

Reviewing accountability of system and legitimizing reforms

When LSAs reveal widespread gaps in student performance, they point to systemic weaknesses, such as ineffective teaching practices, outdated curricula, or lack of resources. This helps policymakers, education authorities, and administrators identify areas of the education system that

need improvement. Governments and education ministries often implement reforms (e.g., changes in curriculum, teaching methods, school management) to improve the education system. LSAs offer a way to evaluate the success of these reforms by comparing performance data before and after their implementation. For example, if a country introduces a new competency-based curriculum, an LSA can assess how well students are performing under this new framework and whether it is improving learning outcomes.

It is well accepted that the ability to read and write, and perform basic operations with numbers, is a necessary foundation and an indispensable prerequisite for all future schooling and lifelong learning [15]. However, in India, various governmental, as well as non-governmental surveys, indicate that there is currently a learning crisis: a large proportion of students currently in elementary school - estimated to be over 5 crore in number - have not attained foundational literacy and numeracy, i.e., the ability to read and comprehend basic text and the ability to carry out basic addition and subtraction with Indian numerals. NEP 2020 states that attaining foundational literacy and numeracy for all children will thus become an urgent national mission, with immediate measures to be taken on many fronts and with clear goals that will be attained in the short term (including that every student will attain foundational literacy and numeracy by Grade 3). The highest priority of the education system will be to achieve universal foundational literacy and numeracy in primary school by 2025.

What are the limits of large-scale assessments?

Narrowing the understanding of education quality

It is important to understand that the purpose of large-scale assessments is to evaluate core competencies in key subjects, primarily through multiple-choice questions, to assess the overall health of the education system. This understanding will enable stakeholders at all levels to implement appropriate interventions, such as curriculum reforms, updates to teaching materials, textbooks, and improvements in teaching practices. Therefore, the quality in education should be understood holistically in terms of educational inputs, educational processes and educational outcomes. However, often LSAs distort the education in several ways [16]-

- LSAs can mislead the public by equating education quality solely with test performance, even though education should involve more than just mastering test-taking skills. LSAs, such as NAS and ASER, are often perceived as indicators of a country's educational quality, triggering media and policy reactions. This focus can shift attention from broader educational goals to test preparation.
- The selection of competencies, subjects, and assessment methods in LSAs is guided by the scale of the assessment, aiming to balance the resources required for conducting the assessment with the value of the data generated. However, this often results in the narrowing of the curriculum, with subjects that are not tested being pushed to the periphery, while those assessed receive more attention and resources.

- LSAs can lead to a focus on "teaching to the test," where instruction is centered on test content and strategies rather than on competency-based learning. Often, excessive time is spent on model tests and test-taking skills at the expense of creativity, problem-solving, and other critical skills.

Underusing and overusing of data

LSAs are systemic level assessments designed to evaluate the overall performance of an education system. These assessments are often conducted with a representative sample of students rather than testing every student in the population. The results obtained from the sample are then analyzed and used to make generalizations about the educational outcomes of the entire student population. The core objective of LSAs is not only to assess student performance but also to explore the relationship between educational outcomes and various background variables such as school characteristics, teaching practices and home environment. By analyzing these correlations, LSAs help identify the factors that contribute to student performance, providing valuable insights into what works and what needs improvement in the education system. Due to lack of this understanding, LSAs data is often not used as desired leading to unintended consequences of testing-

- LSA data is sometimes overemphasized in global, national, or state rankings, leading to a competitive and high-pressure environment. This can cause undue stress on teachers, students, and schools, distracting from a broader, more holistic approach to education.
- LSA data provides a broad overview of education quality in a general manner; it does not capture specific data related to an individual student, teacher, or school. However, it is often observed that individual teachers are penalized and constantly scrutinized by school administrators and district or block-level educational officers.
- Although background variables are collected alongside academic variables, they are often not used to create a comprehensive understanding of the current situation. Merely reporting student outcomes fails to fulfil the true purpose of conducting LSAs.
- There are instances when LSA data is not effectively disseminated to the relevant stakeholders through appropriate channels or reporting methods. This lack of proper dissemination can limit the ability of educators, policymakers, and other stakeholders to fully understand the findings and use them to inform decisions. Often, the data may be presented in a highly technical or inaccessible format that makes it difficult for non-experts to interpret.

Concluding thoughts

A shared understanding of the actual purpose of LSAs among all stakeholders is crucial for their effective implementation and for achieving meaningful improvements in education. This collective understanding fosters a conducive environment where all stakeholders are aligned toward a common goal: continuous improvement in education quality. A culture of evidence-based decision-making emerges, where actions are guided by reliable data rather than assumptions or external pressures. Regular access to accurate, consistent, and reliable data from LSAs is essential for monitoring competencies at each grade level. By tracking students' performance over time, educators and policymakers can identify trends, strengths, and areas that need improvement. This continuous monitoring allows for timely interventions and ensures that teaching methods, curricula,

and resources are aligned with students' needs. The longitudinal data provided by LSAs offers valuable insights into how education systems are evolving and whether the changes implemented are having the desired impact.

In essence, when LSAs are understood and utilized appropriately, they become an indispensable tool for driving positive change in education, ensuring that reforms are evidence-based, targeted, and effective in improving both educational outcomes and overall system performance.

References

1. Brink, S. (2020). Effective use of large-scale learning assessments to improve learning outcomes and education system reform.
2. Banerjee, S. (1993). Revisiting the national literacy mission. *Economic and Political Weekly*, 1274-1278.
3. King, K. (1993). Education for all: national and international implications. *Compare*, 23(1), 5-8.
4. Kumar, K., Priyam, M., & Saxena, S. (2001). Looking beyond the smokescreen: DPEP and primary education in India. *Economic and Political Weekly*, 560-568.
5. Ayyar, R. V. (2005). What Lessons Can DPEP Offer? *Journal of educational planning and administration*, 19(1), 49-65.
6. World Health Organization. (2004). Millennium development goals (No. SEA-HSD-271). WHO Regional Office for South-East Asia.
7. Ministry of Human Resource Development. Sarva Shiksha Abhiyaan. Accessed from: https://dsel.education.gov.in/sites/default/files/2019-05/Manual_Planning_and_Appraisal.pdf
8. NAS. Accessed from: <https://nas.gov.in/>
9. ASER centre. Accessed from: <https://asercentre.org/history/>
10. United Nations. Sustainable Development Goals. Accessed from: <https://sdgs.un.org/goals>
11. Ministry of Education. NCERT. Parakh. Accessed from: <https://ncert.nic.in/parakh/about.php>
12. Azim Premji University. Loss of Learning during the pandemic. 2021. Accessed from: <https://azimpremjiuniversity.edu.in/publications/2021/report/learning-loss-during-pandemic>
13. TaRL in Karnataka. Accessed from: <https://teachingattherightlevel.org/tarl-in-karnataka/>
14. Deccan Herald. India's dipping learning levels. 2022. Accessed from: <https://www.deccanherald.com/opinion/indias-dipping-learning-levels-1124732.html>
15. Ministry of Education. National Education Policy (NEP) 2020. Accessed from: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf
16. UNESCO. (2019). *The promise of large-scale learning assessments*. UNESCO.

ⁱ This article is authored by Shilpi Banerjee. She works as Associate Professor in the School of Continuing Education at Azim Premji University. She can be reached at shilpi.banerjee@azimpremjiuniversity.edu.in. This article can be cited as-

The strengths and boundaries of large-scale assessments, Assessment resources, 2024, Azim Premji University