

---

# Rethinking K-12 Assessment Framework

*Ideas for sound agency design*

Ritika Shah

New Delhi, India  
2020



## Contents

<b>The debate on learning outcomes: what can it tell us?</b>	<b>3</b>
How have learning outcomes been monitored in India historically? . . . . .	3
We have data on learning outcomes but the data is not being utilised as envisioned . . . . .	5
<b>Types and Uses of Assessments?</b>	<b>6</b>
<b>Use 1: Role of learning outcomes assessment in enhancing school accountability</b>	<b>7</b>
What measures of school evaluation for accountability do we have today? . . . . .	7
School accountability to authorities: Compliance with predominantly input standards . . .	7
School accountability to parents: Information black hole . . . . .	8
Can information on learning outcomes assessment be used to hold schools accountable? . . . . .	8
What kind of information works best? . . . . .	9
How can information be reported? . . . . .	9
Can outcomes be linked to explicit rewards and punishments for schools? . . . . .	10
But, will assessments create incentives for schools to manipulate results? . . . . .	10
So, where do we go from here? . . . . .	10
1. Develop assessments for school evaluation . . . . .	10
2. Create a school database . . . . .	11
3. Develop a communication system: Reporting principles and protocols . . . . .	11
<b>Use 2: Role of assessments in evaluating system health and M&amp;E of government programs</b>	<b>13</b>
What is the state of M&E of state education programs in India? . . . . .	13
Limited government research on impact for each rupee spent . . . . .	13
The apex body for quality improvement and research fails to define measurable metrics . .	13
The Department has shown limited capability to monitor the quality of its programs . . .	14
Can information on learning outcomes be used to evaluate efficiency and effectiveness of government programs? . . . . .	14
What are the different parts to M&E? . . . . .	15
How is performance information used in M&E? . . . . .	15
How are results from learning outcomes assessments used in M&E? . . . . .	15
So, where do we go from here? . . . . .	15
1. Define clear performance metrics for the system and measure performance against those	15
2. Per child funding to strengthen monitoring and evaluation . . . . .	16
3. Evidence based interventions . . . . .	16
4. Optimise accessibility and use of system data . . . . .	16
<b>Conclusion</b>	<b>17</b>
<b>Bibliography</b>	<b>18</b>

## The debate on learning outcomes: where are we?

There are inherent information asymmetry problems in education that are characterised using the lens of principal-agent framework. Principal-agent problems exist when the principal hires an agent to act on her behalf but the interests of the principal and the agent are not perfectly aligned. This necessitates that the principal has information to monitor the agent's effort. Typically, in education, there is information asymmetry between parent-child, teacher-child, parent-school/teacher, parent-administrator, administrator-administrator, and administrator-teacher (Figure 1). Administrator means administrator at different units—school, block, district, state and nation (Bergbauer, Hanushek, and Woessmann 2018).

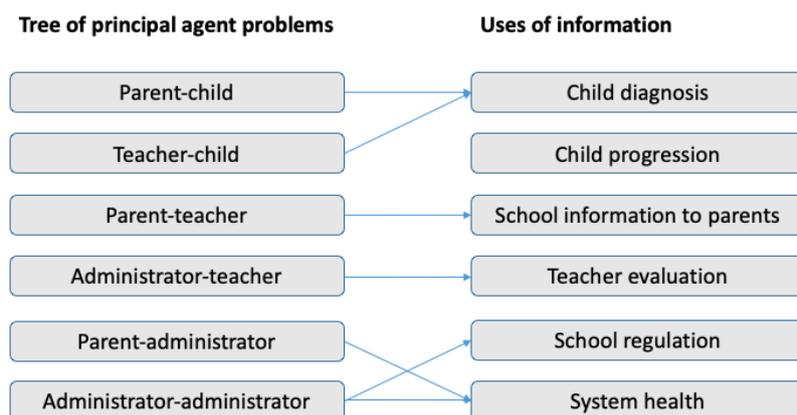


Figure 1: Tree of principal agent problems and different information uses; adapted from Bergbauer, Hanushek, and Woessmann (2018)

These principal-agent problems can be linked to different uses such as child diagnosis, child progression into university, school information to parents, school regulation, teacher evaluation and so on.

A common solution to some of these information gaps is to use learning outcomes assessment. This brief looks at three uses of information, i.e., school information to parents, school regulation, and system health, and explores the role of assessments in each.

### How have learning outcomes been monitored in India historically?

In the two decades preceding 2020, India has had some measures of learning outcomes. These include initiatives by the government institute, National Council of Educational Research and Training (NCERT); non-governmental organisation, Pratham; state governments; and private organisations such as Educational Initiative (Box 1). In addition to national assessments, Tamil Nadu and Himachal Pradesh also participated in the OECD's Programme for International Student Assessment (PISA) in 2009. Besides these, we also have in-class assessments.

**National and state-level achievement tests to assess system health:** With the launch of Sarva Shiksha Abhiyan (SSA) in 2001—India's flagship programme to achieve targets under the global Education for All movement—and an increase in expenditure on education, there was pressure to assess system outcomes. The NCERT was appointed by the Ministry of Human Resource Development (MHRD) to conduct baseline, midline and terminal assessments. Although the NCERT carried out baseline and midline surveys, comparisons across baseline and midline surveys was not feasible. After 2008, NCERT revised NAS methodology allowing for comparison. However, the lack of private schools in the sample and limited analysis of findings to identify action steps has resulted in sluggish adoption of NAS results at the ground level (Oza and Bethell 2013).

In addition, there have been several attempts at the state level, including in Andhra Pradesh, Assam, Delhi, Gujarat, Karnataka and Orissa. Some of these have been criticised for poor quality of methodology, execution and reporting of results (Oza and Bethell 2013).

In the 2000s, the country also saw the emergence of the Annual Status of Education Report (ASER), by Pratham. Since 2005, the report sheds light on the reading and arithmetic levels of children in the 6-14 years age bracket from rural areas every year. The availability of comparable annual data on reading and arithmetic skills of India's children has led to an increased concern and focus on learning levels in policy conversations (Dutt, Kwauk, and Robinson 2016).

**Continuous and comprehensive evaluation through in class tests:** Besides large-scale tests, there are also in-class tests administered in schools predominantly for performance evaluation of students. Section 29(h) of the Right of Children to Free and Compulsory Education Act (RTE) mandates Continuous and Comprehensive Evaluation (CCE) of students of all elementary schools. Under this, formative tests are administered throughout a school year and a final exam is carried out at the end. The scores of formative and summative tests are combined to calculate the end of year grade (Yagnamurthy 2017).

**The Right of Children to Free and Compulsory Education Act did not create a framework to monitor progress on learning outcomes:** The RTE Act did not allude to expected learning outcomes. To correct this, in an amendment to the Act in 2017, states were asked to include class-wise and subject-wise learning outcomes for all grades and alter assessment systems to align with the defined learning outcomes. However, there are currently no ways to test whether these outcomes are achieved by all students.

Separately, in a 2019 amendment, the no-detention policy was revoked to allow states to detain students, if necessary, in standard 5 and 8. However, defining learning outcomes and allowing detention did not help identify or correct reasons for low learning outcomes.

#### **Box 1: Different school and student assessments in India**

We have the following recurrent assessments to measure performance today:

1. **Annual Status of Education Report:** Since 2005, ASER conducts nationwide annual assessments of foundational skills in reading and math for children aged 6-14 in rural areas. ASER is a household rather than school based survey. It samples 600 households in each district—30 villages in each district and 20 households in each village—allowing for a sample of 3,00,000 households across the country.
2. **National Achievement Survey:** NAS uses district as a sampling unit for assessment of grade level competencies for grades 3, 5, and 8. For grade 3 & 5, it measures performance in English, math, and EVS. For grade 8, it also includes social science. It is a school-based survey, however, only looks at government and government aided schools. NAS covers all states and UTs and uses Probability Proportional to Size (PPS) to select districts and schools within districts. Students within schools are randomly selected.
3. **Board exams:** In India, board exams are administered at the end of grade 10 & 12. These are high stakes assessments for the students—grade 10 examinations are typically used to decide whether the student is eligible for the preferred stream of subjects whereas grade 12 decides entry into universities. Boards include the Central Board of Secondary Education (CBSE) and the Indian Council of School Education (ICSE) in addition to multiple State Boards.
4. **In-class tests:** Section 29(h) of the RTE Act introduced Comprehensive and Continuous Evaluation (CCE). CCE replaced one final exam with a combination of formative and summative assessments that are combined to grade a student. These tests are designed and scored by the school. Besides these, there are attempts by state governments to assess, and private organisations such as Educational Initiatives but these are not necessarily recurring.

Although the proliferation of multiple assessment organisations and tools is commendable, the assessments have had a limited impact on driving change in school or administrative practices. There is a lack of systematic thinking on the question of the right tool and right use. This is evident by the low learning outcomes across the country, despite years of assessments pointing in the same direction. This requires a rethink of the assessment framework—what are the different tools and tests available currently, how may they be best used to fill the existing information gaps and what gaps remain to be filled with new tools.

## We have data on learning outcomes but the data is not being utilised as envisioned

Assessments that solve one kind of information problem may not be suitable for another. Tests that evaluate children for progression to university, such as board exams in India, are different from tests that evaluate schools or tests that evaluate teacher performance.

For example, NCERT 2018b regretfully points that despite sharing of NAS results with states/UTs, the results have not been used in ‘education policy, planning and teacher professional development’. This may be possibly because of conflation of multiple uses into one test. NCERT 2018b includes a host of expectations from NAS including ‘attainment of learning outcomes’, ‘identify gaps in the learning’, ‘teachers’ professional development’, ‘provide quality interventions’, and ‘support classroom interventions’. One of the outcomes envisaged from NAS 2017 cycle is the ‘participation of parents and members of school management committee in discussions, planning and monitoring of activities to be undertaken for enhancement of learning outcomes.’ While this is a worthy outcome, it is not possible to use NAS results to drive school or community-level discussions to monitor activities (Centre for Civil Society 2019b).

The conflation of multiple uses into one can also be seen for other assessments such as Board exams. Board exams are designed to assess child performance for progression but have been at times used as indicators of school performance. Unless we isolate the child effort and background from school effort, board exam results are not valid indicators of school effort.

In light of these, it is critical to identify clear and plausible objectives of each test that we carry out currently. What are the principal agent problems NAS can solve? What are the principal agent problems board exams solve? What is the role of in-class tests? How can census assessments be used for school accountability? To answer these, we need to tackle each information asymmetry problem systematically—parent-child, parent-teacher, teacher-child, parent-administrator, administrator-teacher.

## Types and Uses of Assessments

Assessment refers to collection of evidence on learning in a systematic way to judge student learning (Harlen et al. 2002). Assessments can vary in terms of purpose (formative vs summative), uses (student, teacher, school, or program evaluation), stakes (high vs low), frequency (weekly, monthly, annually and so on), scale (in class or cross national), sample (universal vs representative), payment (parents, schools or state) and administration (private vs public tests).

An important caution in any testing practice is to use the test only for the intended purposes. If a single test is used for conflating purposes, the validity of a test is compromised. For this reason, describing the nature and purpose of assessment, at the start, is critical.

The literature on assessment typically divides assessment into two types (OECD 2013)—formative and summative. While this is a useful categorisation, tests should be employed using a problem centric lens—that is, identify information gap that it intends to fill, deploy it only for that purpose with caution for measurement errors. Below we list the different uses of summative and formative assessments:

1. **Summative assessments** indicate the end of course achievements. These are typically high stakes and are used to record or mark achievement, of the student or the school.
  - (a) **School quality evaluation:** Outcomes of summative assessment can be used to inform students and parents the relative effectiveness of a school in imparting learning.
  - (b) **System health:** Outcomes may be used to monitor outcomes of different groups to evaluate the extent to which goals are realised, and to ensure that public funds are used effectively in the process.
  - (c) **Certification or categorisation of students according to learning levels:** Assessments may be used for upward mobility, either from one grade to the next, or from school to university and further.
2. **Formative assessments** are typically low-stake assessments that teachers may use to evaluate a student's progress toward the lesson, unit, or curriculum. It may be used to:
  - (a) **Provide feedback to students and teachers** so that they can make adjustments to their learning process. Such assessments help diagnose gaps allowing teachers to modify content and pedagogy suit students' learning level.

All of these functions can be subsumed under a larger goal of improving learning in schools. In light of the proliferation of assessments and the vexing question of right tool and right use, this brief lays deals with different uses of assessments and learning outcomes in school education.

## Use 1: Role of learning outcomes assessment in enhancing school accountability

School evaluation may be carried out for two reasons, school Improvement or accountability. School evaluation with the objective of school improvement is formative and results in feedback and support to school to raise standards. On the other hand, accountability measures are summative and result in sanctions or rewards.

Accountability relates to providing information to different stakeholders to assess value for money, compliance with standards, and quality of learning and environment (Faubert 2009). It operates at three levels: externally towards state authorities, horizontally towards parents and the larger community, and finally, internal accountability towards meeting the needs of teachers. These may co-exist and be assessed in different ways.

A school evaluation includes four areas: outcomes (academic, personal and social development), classroom practices (quality of learning and teaching), school practices (school as a place of learning and development), and school environment (relations between schools and parents). Within these areas, the focus might be on inputs, processes, or outcomes. Processes are practices that convert inputs into final outcomes. In this brief, we locate the role of outcomes assessments in school accountability to authorities and parents. Some of the questions answered include:

1. What measure of school accountability do we have today? What is the role of outcomes assessment in it?
2. How is accountability to public authorities different from accountability to parents?
3. Can information on learning outcomes assessment be used to hold schools accountable? What kind of information works best? Can outcomes be linked to sanctions or rewards for schools?
4. What are some of the design considerations to keep in mind while developing an assessment framework?

### What measures of school evaluation do we have today?

#### School accountability to authorities: Compliance with predominantly input standards

The Right to Free and Compulsory Education Act (RTE) 2009 is the central law that governs school education in India. The Act lists the standards that a school has to comply with. All the standards relate to inputs. These standards include: the number of teachers, building, minimum number of working days and hours per teacher per week and in a year, learning equipment, library, play material, games and sports equipment (Schedule 9, RTE Act).

Section 18 of the Act also makes recognition of schools mandatory; any school that does not comply with the stated standards can be penalised or be shut down. While the norms and standards also apply to government schools, the provision to penalise or shut schools down does not. In this, the Act discriminates between private and government schools.

There is also the [School Education Quality Index](#) developed by the Niti Aayog. However, it relies on state and UT as the unit of comparison and not the school.

In addition to the RTE Act, schools are also governed by state rules and regulations.

To assess adherence with all the stated norms, the education departments rely on school inspections. As noted in the case of Delhi (Centre for Civil Society 2019b), an inspector may, in addition to compliance with input parameters, assess education quality through teacher and class/homework observations. There is, however, no objective, reliable, valid, or comparable information on outcomes on which schools are evaluated.

## School accountability to parents: Information black hole

Using the principal-agent theory, we understand that parents (principals) delegate the responsibility of teaching their child to teachers (agents). Parents, however, cannot directly observe or monitor the effort a teacher puts in to ensure that their child is learning. While schools have internal tests designed and scored by the teachers, they are aligned with conceptions and instruction of teachers. Further, they do not allow the parent to separate the effect of child effort from teacher effort. It is this information asymmetry that necessitates establishment of horizontal accountability of schools to parents. If parents had complete information about teacher effort, they could maximise child education through investing in the best schools and teachers.

In other countries, this is typically done either through sharing inspection reports and school test results online. Models of public reporting vary. It can either be a ranking of schools (example, league tables in the UK) or a portal (example, [Myschool](#) portal in Australia). The portal provides school-wise information across different parameters including but not limited to performance.

### What does [Draft National Education Policy 2019](#) offer on school accountability?

Below is a quick recap of how the draft National Education Policy aims to bridge the gaps highlighted above:

- 1. Increased use of public disclosure as a tool for oversight and accountability.** Public disclosure on the school and SSRA website will include information on number of classrooms, students, teachers, subjects, fees, and overall student outcomes on standardised evaluations such as NAS and SAS.
- 2. Learning outcomes as a tool for transparency but not regulation:** Suggestion to publicly disclose “finances, procedures, and educational outcomes.” Outcomes will be used to inform parents, irrespective of school type—public or private.
- 3. Regulation on “few basic parameters”:** Such as safety, security, basic infrastructure, the number of teachers across subjects and grades, probity, and sound processes of governance.

The Australian portal only compares performance to schools that have students with similar backgrounds and students who were at the same learning level in the beginning, ensuring a fair comparison of schools with different resources (Centre for Civil Society 2019b).

In India, there are no central rules that require the inspecting officials to share the inspection reports online. At present, there is negligible research on state variance in inspection methods and reporting practices. Centre for Civil Society (2019a) highlighted that, in Delhi, inspection reports are not available online.

State recognition of the school is the only marker of school quality. However, research dating back to 2002, shows that the recognition norms are not applied uniformly and the quality of recognised varies widely (De, Noronha, and Samson 2002). In addition, many parents have opted for low-fee unrecognised schools that stand out of government scrutiny.

In sum, the lack of government norms enabling school transparency, objective information on outcomes, non-reporting of inspection results, and the varied application of norms compound the information gap parents face in holding schools accountable.

## Can information on learning outcomes assessment be used to hold schools accountable?

The use of learning outcomes to evaluate schools' performance, included under the umbrella of standards based reforms, is increasingly used globally as a tool for school accountability. The objective of standards based reforms is to have clear, measurables and achievable goals for schools to work toward. The extent and degree of such standards-based reforms varies globally. There are different parts to the problem:

## What kind of information works best?

In a recent paper, Bergbauer, Hanushek, and Woessmann (2018) study how different types of information affect student outcomes. The authors, using PISA data of 59 countries over six waves between 2000 and 2015, organise testing into four types:

1. Internal testing includes teacher generated-tests: Results of these tests are difficult to compare across teachers (and schools) and thus, do not solve the parent-teacher principal-agent problems. If not standardised, parents also can't compare information across schools
2. Standardized testing with external comparison includes student-based tests (centralized exit exams) and school-based tests (for example, under No Child Left Behind)
3. In between lie standardized testing without external comparison lie internal tests that may be standardised but do not provide for comparison to other schools or students
4. Internal teacher monitoring done through inspectors

They find that only the expansion of standardised tests with external comparisons, school and student-based, is associated with improvement in student achievement in international tests. The effect of these tests is higher in low-performing countries. The authors find that internal testing and internal teacher monitoring do not impact student achievement on international tests (source).

In India, we are currently using internal testing and internal teacher monitoring. Given the low performance on different assessments, and in light of these findings, there is an urgent need to escalate to standardised testing that allows for external comparisons at school level.

## How can information be reported?

In an experiment in Rajasthan, Afridi, Barooah, and Somanathan (2017) provided schools and/or parents with information on performance of students on standardised tests. There were four kinds of report cards—first two shared with parents and the latter with schools:

1. Report card for parents (1): student's *absolute performance* in Hindi, English and Math as well as her *relative performance* by ranking her in her grade
2. Report card for parents (2): This included all of the elements of first. In addition, it included the relative performance of *all students and schools in the panchayat* based on combined scores, for the relevant grade
3. Report card for schools (1): Reported the average, subject specific score for *each grade in the school* and the proportion of students at different levels of competence in reading, writing and numeracy in each grade at the school level
4. Report card for schools (2) showed the grade-averaged score in each subject of *all schools* in the panchayat

These four report cards were in the following manner to note how different information impacts stakeholders:

1. Group 1: Only report card for parents (1) was shared with parents
2. Group 2: Report card for parents (1) was shared with parents and report card for schools (1) was shared with schools
3. Group 3: Report card for schools 1 & 2 were shared with schools but parents had only had the first report card for parents
4. Group 4: Full information was given to parents and schools. Report card for parents 1 & 2 was shared with parents and report card for schools 1 & 2 was shared with schools

Findings revealed that significant improvements were observed only in treatment group 4 when parents were informed of their child's performance and of all others in the Panchayat and when schools were informed of their students' performance and performance of other schools. In contrast, there was no significant improvement if only schools were informed. In addition, public schools did not see an improvement in learning outcomes. In sum, the authors found that dissemination of information to both parents and schools, on relative and absolute quality of schools, is a low-cost and effective tool for improvement in learning outcomes, especially in private markets where there is threat of competition.

### Can outcomes be linked to explicit rewards and punishments for schools?

Learning outcomes can be used to affect school accountability through direct government action or community action. Direct measures include use of rewards and sanctions such as increased funding, penalty, and award or withdrawal of recognition. Public reporting of information to guide parental choice is an example of indirect use.

There is evidence, from the United States, that measures such as monetary awards, threat of government take over of non-performing schools and offering vouchers to students of non-performing schools result in improved performance. The gain in performance is higher than that achieved by only reporting outcomes (Hanushek 2005).

These systems are, however, built and tested over time. Long before the United States implemented No Child Left Behind (NCLB), a policy that held schools accountable for student achievement, individual states had developed their own testing methods and accountability systems. In addition, NCLB was only applicable to public schools that received federal funds. Given that information on outcomes can be misused, misinterpreted, or inappropriately applied, it is critical that the system reaches maturity in method and application of assessment results before attaching results to rewards or sanctions to individual schools.

### But, will assessments create incentives for schools to manipulate results?

Use of assessment for regulation or for public disclosure impacts a school's sustainability. While poor performance can be a reason for shut down where outcomes are used for regulation, even public disclosure can impact school enrolment and its sustainability. Given the stakes attached, schools may engage in fraudulent and manipulate results. This can be done through classification of certain students as mobile/disabled to avoid testing of select students, admission of students based on demonstration of learning abilities or through cheating on such tests.

Adequate measures must be taken to prevent fraudulent practices in areas where there is a high threat of manipulation to ensure high result fidelity. To prevent selection of better performing students, schools should be prevented from engaging in test based admissions, allowing all students, irrespective of learning achievement, an equal opportunity for enrolment in the school of their choice. The criteria for inclusion and exclusion of students, or reclassification, should be clearly defined and strictly enforced such that incentives to manipulate are minimised. For prevention of cheating during tests, the testing agency can either deploy external supervisors or conduct controlled retesting of students.

### So, where do we go from here?

#### 1. Develop assessments for school evaluation

There are several design considerations in the execution and reporting of test results. Below we list few of these:

**Test different models of testing:** A critical consideration in developing assessments for school accountability is to test different models that allow for optimal use of limited state capacity. What is the role of the state in enabling census assessments? Should the state directly assess or enable existing or new civil society organisations to play a role? Can boards be a channel for higher frequency or key stake testing and reporting? What is the role of private providers of tests? While in some countries, census assessments are led by the government, there are regions that have explored other models. For instance, Alaska, in its [state regulations](#), in the US and Ireland (Claire, Deborah, Paul, et al. 2014)

require schools to choose from private providers of standardised tests and report results to concerned authorities. Separately, in northern Ireland, schools choose standardised tests offered by Awarding Organisations.

**Status vs value-add measures:** There are two ways in which school performance can be reported: status and value-add measures. A status measure indicates student performance against set standards whereas a value-add measures indicate a student's performance against past performance or in other words, the year to year growth in achievement.

The two measures incentivise different actions by the school. The status measure, for example, has the advantage of pushing school to ensure that all students, including the low performing students, achieve minimum proficiency level. Value-add measures, on the other hand, are considered as the fairer of the two as they take into consideration the level each student starts from. This means that schools do not have to necessarily achieve a minimum for all as long as targets for growth are met.

Value add measures require additional data points such as annual tests to calculate progress and assignment of identification code to each child to link results annually and over the years.

Status measures are considered transparent, easier to understand by the public, and may potentially allow for a greater dialogue between parents and schools, especially if public disclosure is the only use.

Depending on the objectives, the government may use status measures, growth measures, or a combination of the two.

**Universal vs partial inclusion:** Universal inclusion raises questions: Is it fair to include students who joined a given school less than a year ago? Can schools be held accountable for students with learning disabilities? There is a tension here: if mobile and disabled students are excluded, schools' incentives to work with them reduces, whereas if they are included, it is unfair to compare school's with large proportions of such children with other schools with relatively fewer mobile and disabled student population.

One way to resolve the tension around mobile students is to only include students who have spent at least a year in the school. In the NCLB scheme in the US, for example, only those students who have spent a year in a school are included.

For disabled students, schools could be offered financial incentives to offset the higher cost involved.

## 2. Create a school database

Currently, there is no comprehensive database of K-12 schools in India. For any policy that upholds attainment of minimum learning outcomes as an objective, without discrimination between students of public and private schools, an essential requirement is to cover all schools and all children.

It is the most vulnerable of the lot that are likely to be enrolled in low-fee schools. Any policy should be designed such that unrecognised schools are encouraged to formalise and not alienated in the process.

## 3. Develop a communication system: Reporting principles and protocols

Currently, some state governments collect information on the staffing, facilities, and learning environment at the school level through inspections. However, this information is retained between the school and the officials without any meaningful transfer to parents. The draft policy emphasises on transparency through disclosure of information on the school website. It is essential that sufficient thought is put into the content, understandability and medium of communication.

The *Principles and protocols for reporting on schooling in Australia (2009)* is a useful document that lays the different elements of a good reporting system. Some of the principles include:

1. **High quality data:** Any publicly reported data should be valid, reliable and contextualised. Validity means it should only be used for the purpose it is intended to be used. Any misuse or misapplication of information for purposes outside the defined objectives may compromise validity and have adverse effects.

2. **Easy to understand:** Data on learning outcomes should be presented such that it allows parents to understand their child and schools' performance against standards and cohort learning level. Interschool comparisons should only be done with supporting contextual information. In addition, there should be supplementary information to help parents make sense of the data on schools.
3. **Contextual information:** In addition to finances, procedures, and educational outcomes as the draft education policy suggests, there should be information on a school's philosophy and method to learning, fee structure, staffing, enrolment profiles, and offering of extra curricular activities. This information should be objective, factual and comparable across school types.
4. **Uphold student privacy:** Information on outcomes should be anonymised maintain privacy of the students.

The government must carefully design reporting principles and protocols and ensure those are complied with at all times.

## Use 2: Role of assessments in evaluating system health and M&E of government programs

Since the 1970s, there is a global trend towards adoption of business management principles to government administration referred to as New Public Management (NPM). The concept spread fast from countries where it originated to other developed and developing countries making it a global phenomenon. This marked a preference for a smaller public sector and a dependence on markets for provision of services. The public sector, during the Thatcher regime in Britain, was diagnosed as overly bureaucratic, wasteful and bloated giving birth to a need for a lean, efficient and effective state, in essence, a state that functioned like a business. The change was also driven by social factors such as demands for an ethical government, decentralisation and an increase in public participation. NPM rests on principles of public choice and principal agent theory among others (Tolofari 2005).

Public choice theory applies the principles of market economics to public administration. It argues that just as market players are driven by self interests, so are government officials and bureaucrats. An extension of this is taxpayers' and citizens' demands for fiscal discipline, value for money and accountability from government officials. Principal agent theory led to agencification and separation of functions of provider from the user (Tolofari 2005).

In education, the adoption of NPM led to privatisation, performance measurement and accountability. In the following section, we look at how we evaluate our education programs, how has NPM impacted education administration globally and the lessons these trends hold for India.

### What is the state of M&E of state education programs in India?

Let's take the example of Sarva Shiksha Abhiyan (SSA). SSA, launched in 2000, was the government's flagship program encompassing "variety of interventions for universal access and retention, bridging of gender and social category gaps in elementary education and improving the quality of learning" (Ministry of Human Resource and Development 2018). SSA was also the primary vehicle for implementing the Right to Free and Compulsory Education Act 2009 (RTE). How much money was invested into the program? Were stated outputs achieved? Was the money spent economically, well, and wisely? What are the monitoring mechanisms?

#### Limited government research on impact for each rupee spent

Research by non-government organisations such as Accountability Initiative and development partners such as the World Bank elucidate these questions. The per-child funding under SSA, calculated by dividing total approved budget by the total number of children in government schools, has been on the rise since 2010 except a dip in 2013-14. It increased from Rs 3913 to 6350 in 2016-17 (source). Separately, a value-for-money (VFM) analysis shows that despite the increase in government expenditure on education over the period 2011 to 2014, student learning levels fell instead of rising, and has thus resulted in a negative VFM. While the returns on quality are negative, investments in SSA have had positive VFM as far as access to schooling is concerned (Kingdon, Sinha, and Kaul 2016)

It is critical to point out, as a commentary on the government's monitoring and evaluation framework, that none of this research is not conducted by the government. State reports offer no clear answers to the questions on effectiveness and efficiency of investments.

#### The apex body for quality improvement and research fails to define measurable metrics

The National Council of Educational Research and Training (NCERT), an autonomous body that categorically assists in qualitative improvement, released a report in 2018 on the quality monitoring mechanisms in states. The objective was to study administrative structure for quality monitoring, the components of monitoring framework, analyse periodicity, frequency and follow-up mechanisms, and finally, the perception of education functionaries on monitoring mechanisms and its user-friendliness. What stands out in the report that evaluates the effectiveness of our quality monitoring mechanisms is the section that defines quality: 'There is no one definition of the term 'quality' as it is a relative

attribute. It generally refers to a high degree of goodness, worth or excellence in an object, process or system. In the social context, it represents rather an abstract idea, which needs to be exemplified through certain indicators. In education, as per common man’s parlance, quality is seen in terms of the input, process and output. Quality elementary education is multidimensional, dynamic and complex. In order to attain a desirable output in terms of the learner’s achievement, both in curricular and co-curricular areas, necessary inputs (planning) and quality processes (implementation) need to be provided’ (NCERT 2018a). This is at best tautological and at worst, meaningless. How do we monitor quality if we don’t define meaningful and measurable metrics?

### The Department has shown limited capability to monitor the quality of its programs

The Ministry of Human Resource Development through NCERT developed a program to measure quality in 2005-06. Tools were developed for different levels, including school, cluster, block, district, and state. Figure 1 lists the different themes covered under quality monitoring tools. While a list of indicators are measured, there are no linkages between resources, inputs, processes, outputs and outcomes. For example, while learner’s assessments is one of the considered indicators, there is no link between learning and teacher preparation or classroom practices.

The department has not established rigorous mechanisms to monitor programs. Data sources through NAS has thus far not been analysed with respect to funding allocations to state. As a consequence, the department is not well placed to make evidenced based decisions. There is little research to indicate what drives costs and what generates results. Weaknesses in current monitoring arrangements, captured by NCERT (2018a) and Bhatta (2016) have undermined the department’s ability to assess progress against established policy objectives.

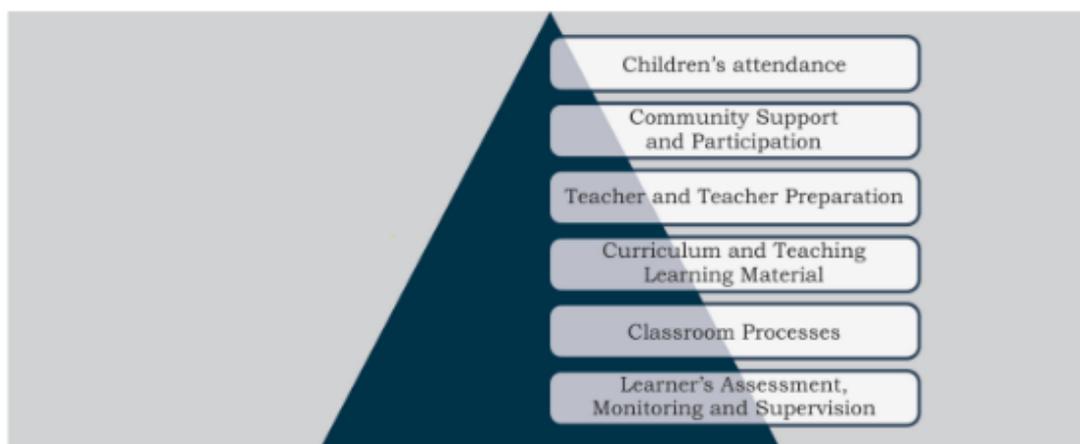


Figure 2: Quality Dimensions covered in NCERT’s Quality Monitoring Tools (NCERT 2018a)

### Can information on learning outcomes be used to evaluate efficiency and effectiveness of government programs?

The adoption of NPM to education administration has led to ‘managerialism’ in education and a rise in powers of school leadership and local community, away from local authorities, in budget, staff appointment, and planning. This is accompanied with increased emphasis on accountability of school to parents and change in fund allocation to per-child basis. In addition, there is a transformation in the role and responsibilities of the head teacher who is now seen as a manager exercising functions such as finance and performance management like a business leader would.

A second consequence is ‘performativity’. Performativity is marked with an increase in the use of standards to hold schools accountable but also a devolution in areas such as hiring, curriculum and pedagogy to allow for diversity. Besides, there is an emphasis on parental choice (Tolofari 2005).

As a result of these movements, monitoring and evaluation (M&E) has become a critical part of program delivery, policy making and public accountability. Monitoring is defined as “ongoing,

systematic collection of information to assess progress towards the achievement of objectives, outcomes and impacts,” whereas evaluation is “the systematic and objective assessment of an ongoing or completed project, programme or policy, its design, implementation and results, with the aim to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability” (*Guidelines for Project and Programme Evaluations 2009*).

### **What are the different parts to M&E?**

An M&E begs the question of whether a government program has achieved its stated objectives in the most economical, efficient and effective manner. There are various parts to an M&E system. Learning outcomes assessments forms one part of it. Other parts may include: teacher evaluation, inspections, a statistical data system, and school level information. Countries such as Chile, Brazil, Singapore and Australia are recognised as having well developed M&E systems (UNESCO 2016).

### **How is performance information used in M&E?**

There is an increase in use of performance information in budget planning and decision making. However, the kind of data and the way in which it is used varies across countries. The use may include stating goals (may or may not include learning outcomes), measuring progress and reporting outcomes. In addition, countries use evaluation results, value for money analyses, and spending reviews to inform decision making. A number of countries have also introduced performance oriented budgeting as part of accountability reforms. This essentially involves defining inputs, processes, outputs, and outcomes and evaluating the links between these. Austria is one such country that introduced performance based budgeting in education to see how inputs and outcomes are linked. While the use of impact-based funding has potential for improved decision making, it is at a nascent stage, especially in OECD countries (OECD 2017b).

### **How are results from learning outcomes assessments used in M&E?**

In particular, results from learning outcome assessments are used for different purposes: informing policy, monitoring standards, refining standards, identifying correlates of achievement, promoting accountability, and informing discussion. The use of outcomes in accountability depends on who is held accountable (student, teacher, school, government officials) (Greaney and Kellaghan 1996).

### **So, where do we go from here?**

In 2018-19, Samagra Shiksha Abhiyan subsumed SSA and Rashtriya Madhyamik Shiksha Abhiyan (RMSA). This means that K-12 education will be treated as one unit without segmentation in pre-primary, primary, upper primary and secondary education.

#### **1. Define clear performance metrics for the system and measure performance against those**

The primary objective of Samagra Shiksha is ‘enhancing learning outcomes’. However, there are no stated goals. For example, all states should achieve xyz level of learning by 2025.

In an amendment to rule 23(2)(c) of the RTE Central Rules 2010, all states have been mandated to prepare ‘class wise, subject wise learning outcomes.’ States have been slow in adopting learning outcomes. There is ambiguity around what these outcomes are and how and if they can be measured (Kapoor 2018). Public officials require clear performance metrics, and systems to measure progress on each. Typically, one finds that in pilots goals are clearly stated and evaluated at the end of the study. This allows different agencies and officials to be aligned towards the goals.

In the latest Performance Grading Index developed by MHRD, states get full score if 75% of students answer NAS questions correctly. While it is helpful to develop an index to foster competitive federalism, we also need clear metrics on what is ‘unacceptable’ vis-a-vis outcomes in education.

Without identifying schools, districts and states that are in urgent need of help with regards to learning outcomes, we are shooting in the dark with interventions.

## **2. Per child funding to strengthen monitoring and evaluation**

The government website indicates that, under Samagra Shiksha Abhiyan, funds will be 'allocated based on an objective criteria based on enrolment of students, committed liabilities, learning outcomes and various performance indicators'. While the government has not yet notified a formula but a transition from the current system of block grants to funding on a per-student basis can increase school accountability. A number of developed countries, including the US and UK, allocate funding based on school enrolments (OECD 2017a). Such a system automatically weeds out inefficient schools as schools that fail to enrol students either due to unsatisfactory performance or lack of demand lose funding and are shut down (Kingdon, Sinha, and Kaul 2016). Currently, some states such as Rajasthan and Jharkhand have integrated small schools into one large school complex for resource sharing and economies of scale. Per child funding model that takes into consideration the enrolment but also learning outcomes of schools as parameters for consolidation can further enhance this process.

## **3. Evidence based interventions**

It is critical to understand what works and what does not and ensure that investments are backed by evidence. Under [Samagra Shiksha Abhiyan](#), the plan is to improve learning outcomes by interventions related to 'teacher and technology'. How have we arrived at this conclusion? If funding decisions are not backed by evidence, or if evidence is flawed, the department will not achieve desired results. The department should identify a set of indicators—realistic outputs and outcomes—that can be measured and use those to evaluate the impact of interventions.

## **4. Optimise accessibility and use of system data**

Currently, the results of national monitoring of public school education are dispersed and are not available in an easy to understand manner. While some information is available but spread across platforms, such as [UDISE](#), [performance grading index](#), and the [MHRD website](#), other information such as school inspection reports and variation in spending per child across schools, districts, and states is not available. Further, no single platform collates this information to summarise performance of different units of evaluation, that is, school, zone, block, district, city, state and nation. While the Union Government has initiated work on a new platform called SE Shagun, it is critical that information is comparable and supported by analysis from different researchers—in house or contracted—to illuminate what the results actually mean (OECD 2012).

If the platform offers local contextual information, against nationally set benchmarks, it is also likely to strengthen local monitoring.

## Conclusion

The latest draft of the [National Education Policy](#) released in 2019, led by Dr Kasturirangan, proposes multiple ideas to reform existing assessment practices and also introduces new ways to employ assessments. The document lays out a number of new test uses—low stakes assessments to personalise teaching, national achievement survey for periodic health check up, census assessment for disclosing school performance and board exam reforms for student progression.

The draft emphasises on increasing the use of low-stakes tests, carried out during the process of teaching-learning, to assess achievement level of each student and develop personalised learning plans accordingly. To reduce the high stakes attached to board exams, and in cognizance of the stress it creates for students, the draft pushes for increased flexibility in subject uptake and the change from the current end of year exam system to allow for multiple attempts for different subjects at the end of a semester.

In addition to the sample-based National Achievement Survey for ‘periodic health check-up’, the document suggests states to introduce a census-based State Assessment Survey. Census assessments will be used only for ‘developmental purposes, by sharing it with teachers, students, and their parents’. At one point, the draft states that the data may be shared with the School Management Committee and the community.

Yet, even as the draft policy lays out these different tests, it lacks lucidity on the the problem each of these tests solve, the information gap it fills, and the use it will be put to. Critical terms such as ‘developmental purposes’ and ‘health check-up’ remain vague and open for interpretation.

The draft fails to answer the fundamental challenge of information asymmetry in education services. How can we systematically address the information asymmetry problems faced by different stakeholders including parents, teachers, school and administration? What are the different kinds of information required by each? What is the role of learning outcomes assessment—in class or large scale? How will the information generated through these will be used to tackle the challenges? It is these questions that this brief calls attention to and provides a starting point for reference.

## Bibliography

- Afridi, Farzana, Bidisha Barooah, and Rohini Somanathan. 2017. "Improving learning outcomes through information provision: Evidence from Indian villages."
- Bergbauer, Annika B, Eric A Hanushek, and Ludger Woessmann. 2018. *Testing*. Technical report. National Bureau of Economic Research.
- Bhatty, Kiran. 2016. *Does the Government's Monitoring of Schools Work? A Study of the Frontline Education Bureaucracy in India*.
- Centre for Civil Society. 2019a. *Anatomy of K-12 Governance in India*. Centre for Civil Society. <https://ccs.in/sites/default/files/Anatomy-of-K-12-Governance-in-India.pdf>.
- . 2019b. *Role of Learning Outcomes in Education Governance*. Centre for Civil Society. <https://ccs.in/sites/default/files/research/draft-blueprint-on-use-of-learning-outcomes-in-education-governance-april-2019.pdf>.
- Claire, Shewbridge, Nusche Deborah, Wright Paul, et al. 2014. *OECD reviews of evaluation and assessment in education: Slovak Republic 2014*. OECD Publishing.
- De, Anuradha, Claire Noronha, and Meera Samson. 2002. "Private schools for less privileged: Some insights from a case study." *Economic and Political Weekly*: 5230–5236.
- Dutt, Shushmita Chatterji, Christina Kwauk, and Jenny Perlman Robinson. 2016. *Pratham's Read India Program*. Centre for Universal Education at Brookings. <https://www.brookings.edu/wp-content/uploads/2016/07/FINAL-Read-India-Case-Study.pdf>.
- Faubert, Violaine. 2009. "School evaluation: Current practices in OECD countries and a literature review."
- Greaney, Vincent, and Thomas Kellaghan. 1996. *Monitoring the learning outcomes of education systems*. The World Bank.
- Guidelines for Project and Programme Evaluations*. 2009. Australian Development Cooperation. <https://www.oecd.org/development/evaluation/dcdndep/47069197.pdf>.
- Hanushek, Eric A. 2005. "Why the federal government should be involved in school accountability." *Journal of Policy Analysis and Management* 24 (1): 168–172.
- Harlen, Wynne, Ruth Deakin Crick, Patricia Broadfoot, Richard Daugherty, John Gardner, Mary James, and Gordon Stobart. 2002. "A systematic review of the impact of summative assessment and tests on students' motivation for learning."
- Kapoor, Taanya. 2018. "Mapping Learning through Outcomes: Understanding the amendment to the RTE." <https://accountabilityindia.in/blog/mapping-learning-through-outcomes-understanding-the-amendment-to-the-rte/>.
- Kingdon, G, S Sinha, and V Kaul. 2016. *Value for money from Public Education Expenditure on Elementary Education in India*. Technical report. Discussion Paper Series, Education Global Practice, South Asia Region, World ...
- Ministry of Human Resource and Development. 2018. "Samagra Shiksha Abhiyan." <https://timesofindia.indiatimes.com/city/raipur/Chhattisgarh-among-bottom-10-in-education-development/articleshow/21155205.cms>.
- NCERT. 2018a. *A Study of the Functioning of Quality Monitoring Mechanisms in States and UTs*. NCERT. <http://www.ncert.nic.in/departments/nie/dee/publication/pdf/Quality-monitoring-mechanism.pdf>.
- . 2018b. *Post NAS Interventions*. NCERT. <https://www.brookings.edu/wp-content/uploads/2016/07/FINAL-Read-India-Case-Study.pdf>.
- OECD. 2012. *Education system evaluation*. doi:<https://doi.org/https://doi.org/10.1787/9789264117020-9-en>. <https://www.oecd-ilibrary.org/content/component/9789264117020-9-en>.

- . 2013. *Student assessment: Putting the learner at the centre*. doi:<https://doi.org/https://doi.org/10.1787/9789264190658-7-en>. <https://www.oecd-ilibrary.org/content/component/9789264190658-7-en>.
- . 2017a. *Distributing school funding*. doi:<https://doi.org/https://doi.org/10.1787/9789264276147-7-en>. <https://www.oecd-ilibrary.org/content/component/9789264276147-7-en>.
- . 2017b. *Evaluating the use of school funding*. doi:<https://doi.org/https://doi.org/10.1787/9789264276147-9-en>. <https://www.oecd-ilibrary.org/content/component/9789264276147-9-en>.
- Oza, Jayshree, and George Bethell. 2013. *Assessing Learning Outcomes: Processes, Policies and Challenges*. Technical report. Department for International Development.
- Principles and protocols for reporting on schooling in Australia*. 2009. Australian Curriculum, Assessment and Reporting Authority. <http://www.educationcouncil.edu.au/site/DefaultSite/filesystem/documents/Reports%20and%20publications/Publications/Measuring%20and%20reporting%20student%20performance/Principles%20and%20protocols%20for%20reporting%20on%20schooling%20in%20Australia.pdf>.
- Tolofari, Sowaribi. 2005. “New public management and education.” *Policy futures in education* 3 (1): 75–89.
- UNESCO. 2016. *Designing effective monitoring and evaluation of education systems for 2030: A global synthesis of policies and practices*. UNESCO Education Sector. <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/pdf/me-report.pdf>.
- Yagnamurthy, Sreekanth. 2017. “Continuous and comprehensive evaluation (CCE): policy and practice at the national level.” *The Curriculum Journal* 28 (3): 421–441.