



# Improved Learning Outcomes for Asia - Improving Learning Outcomes for the Philippines

## **Strengthening performance management: Case studies and key insights**

Submission Date: 29/01/2024

Prepared by Delivery Associates

Submitted by: **RTI International**  
3040 East Cornwallis Road  
Research Triangle Park, NC 27709-0155  
Tel: (919) 541-6000

This document was produced for review by the Philippines Department of Education and the United States Agency for International Development



# TABLE OF CONTENTS

	Page
Abbreviations and Acronyms .....	2
1 Overview and Scope.....	3
2 Introduction .....	6
3 Summary and key insights .....	8
4 Case studies .....	12
4.1 Indonesia (Education) – <i>System with limited performance and accountability levers</i> .....	12
4.2 New South Wales (NSW) (Education) - <i>System with mixed performance and accountability levers</i> .....	20
4.3 Punjab, Pakistan (Education) - <i>System with strong performance and accountability levers</i> .....	27
4.4 New South Wales (Health) - <i>System with mixed performance and accountability levers</i> .....	37
5 Sources .....	I

# ABBREVIATIONS AND ACRONYMS

AEO	Administrative Education Officers
AKM	<i>Asesmen Kompetensi Minimum</i> (Minimum Competency Assessment)
BPS	Badan Pusat Statistik (Central Bureau of Statistics)
CEO	Chief Executive Officer
DCO	District Coordination Officer
DDEO	Deputy District Education Officers
DEO	District Education Officers
DFC	District Field Coordinators
D-IV	<i>Diploma Empat</i>
DRC	District Review Committee
EAV	Emergency Access View
EDCOM II	The Second Congressional Commission on Education
ILOA	Improving Learning Outcomes for Asia
ILO-PH	Improving Learning Outcomes for the Philippines
KPI	Key Performance Indicators
LHD	Local Health Districts
LND	Literacy and Numeracy Drive
MEA	Monitoring and Evaluation Assistants
MI	<i>Madrasah Ibtidaiyah</i>
MSS	Minimum Service Standards
MT	<i>Madrasah Tsanawiyah</i>
NAPLAN	National Assessment Program Literacy and Numeracy
NSW	New South Wales
PMIU	Program Monitoring and Implementation Unit
RTI International	Research Triangle Institute International
S1	<i>Strata-1</i>
SD	<i>Sekolah Dasar</i> (Primary School)
SLO	Student Learning Outcomes
SMART	Specific, Measurable, Achievable, Realistic, Timebound
SMP	<i>Sekolah Menengah Pertama</i> (Junior Secondary School)
SOP	Standard Operating Procedure
SMU	Spatial Monitoring Unit
USAID	United States Agency for International Development

# I OVERVIEW AND SCOPE

EDCOM II has been tasked with driving Education Reform in the Philippines to improve learning outcomes for all children. Doing so will ensure that Filipinos can not only secure their own future but can significantly contribute to the growth of the Philippines and the region as a whole. Of EDCOM II's 28 priorities, a number have been identified for immediate focus, including Priority 25: *'Integrated performance management and accountability system'* which aims to strengthen the Education system's ability to manage performance and strengthen accountability around learning outcomes at teacher, regional and system levels.

Despite the implementation of various practices and frameworks for performance management, such as the Philippine Development Plan (PDP), the Results-Based Performance Management System (RPMS), and the Performance-based Bonus scheme (PBB), there is limited success in driving performance improvements in the system.

One of the main concerns is that performance management and accountability frameworks at the system, organizational, and individual levels are isolated and do not function in cohesion as part of a holistic performance management and accountability system that drives improvements in student learning outcomes. As an example, there is a weak relationship between the Individual Performance Commitment and Review Form (IPCRF) for school teachers, and the Office Performance Commitment and Review (OPCR) of school principals.

At various levels of the system, there are additional gaps identified in discussions and focus groups with officials and technical personnel from DepEd, CHED, and TESDA, including<sup>1</sup>:

- Focus on compliance, reporting and the collection of means of verification within performance management and accountability systems, over evaluating practices and strategies
- Emphasis on processes rather than outcomes in performance management tools, for example for school teachers there is more focus on teaching procedures for teachers rather than potential learning outcomes
- Limited incentives for improving performance on outcomes, including student learning outcomes
- Limited focus on reflection and the diagnosis of performance gaps
- Insufficient time for actual coaching and mentoring at the school level
- Absence of a feedback culture in the system

These gaps collectively underscore the need for a more integrated and outcome-focused approach to performance management and accountability within the education sector.

---

<sup>1</sup> Source: Miseducation - The Failed System of Philippine Education, EDCOM II Year 1 report

ILO-PH has the opportunity to draw on the collective expertise and experience of Delivery Associates, RTI International, and the ILOA project to conduct a *Performance and Accountability Benchmarking Review* to identify practices in performance and accountability across large-scale and complex systems (both within and outside education) focusing on system, regional and frontline levels. Drawing on experiences elsewhere will inform the next phase of this engagement: a detailed review of the current state of performance and accountability in the Philippines education system, assessment of the current system against the identified best practices, and the identification of specific recommendations for its strengthening.

To identify global practices of embedding a culture of performance and accountability in complex systems which can serve as a guide on addressing the challenges identified in the Philippines, this report explores the following key questions:

- In education sectors, is accountability and performance management at teacher, school, regional and system levels centered around learning outcomes and what are the characteristics that have made these effective/ineffective?
- In education and non-education sectors, how are targets for accountability and performance management cascaded throughout a system i.e. at organizational, managerial, and frontline staff levels?
- In education and non-education sectors, what mechanisms and tools are used for effective accountability and performance management?

To answer these questions, we reviewed performance management and accountability practices across **three education systems: Indonesia, New South Wales (NSW), Australia, and Punjab, Pakistan**. These systems were chosen to represent both centralized structures (Punjab, NSW) and decentralized structures (Indonesia).<sup>2</sup> The selection also allows us to gather insights from a maturer system (NSW Education) as well as developing systems (Punjab, Indonesia). Punjab and NSW are chosen for this analysis since in both these geographies, ambitious education reform has been undertaken in the last decade, and in both instances performance management and accountability have been key drivers of the reform efforts. In addition, Indonesia offers an example of a Southeast Asia country with similar challenges as the Philippines, and which in recent years has begun to implement and experiment with new practices to enhance performance of the education sector.

The report also examines the **health sector in New South Wales (NSW), Australia**. Despite the differences between health and education systems, this presents an opportunity to understand the application of common performance management and accountability practices in a different context. NSW Health offers an insightful case as, like NSW Education, the Health sector has set an ambitious agenda over the past decade and has adopted many effective performance management and accountability practices.

---

<sup>2</sup> Pakistan has a decentralized education system at the national level. Within individual provinces, such as Punjab, education is centralized under the provincial government.

In our case studies, we examine how performance management and accountability practices operate collectively across various levels—teacher, school, regional, and system-level—to drive a system towards a high-performance culture.

## 2 INTRODUCTION

Taken together, the case studies in this Report highlight the core components of an effective performance and accountability framework – a comprehensive system to optimize individual and system-wide performance through establishing outcome goals, tracking, and incentivizing progress against them. The case studies show that effectiveness is not determined by implementing a set of individual policies, tools, and structures, but rather how they are collectively used to drive a **culture** that empowers its people to achieve a system's ambitious agenda.

The components of a performance and accountability framework that are featured in these case studies include:

- **System leadership** that prioritizes a culture of high-performance around targets.
- **Outcome focused targets** and **supporting indicators** that are well understood and communicated throughout the system, to define and quantify a system's aspiration.
- **Quantitative data** which is **available** and **regularly shared** across a system to gauge system performance, and to identify areas for additional focus or support.
- **Targeted and tailored support** for system actors to help them improve performance, including a focus on specialized support for underperformers, driven by an understanding of data and evidence.
- **Performance routines and reporting** with key system actors focused on reviewing progress, problem solving, and decision making to unlock barriers to improvement.
- **System engagement** and site-visits/fieldwork that surfaces key challenges and opportunities to drive improvement.
- **Public engagement** that keeps the public up to date on progress.

Some of these components may require adaptation to different parts of the system. For example, targets may feature in performance agreements for senior leadership roles, but not for frontline staff who may be engaged in a more indirect way around targets (for example, communication of the reform with a focus on enhancing practice and improved outcomes). Similarly, formal performance routines that focus on a broader strategic reform are effective at driving a performance culture at senior levels, but not so for frontline staff who are rightly focused on teaching practice in their classroom. To be effective, embedding a performance and accountability culture across a system requires a nuanced and tailored application of these components to frontline staff, middle management, and senior levels.



The table below draws out select examples of the application of these components as demonstrated in the featured case studies:

Component	Fails to embed a performance and accountability culture	Embeds a performance and accountability culture
<b>System leadership</b>	Singular focus on performance, and overly reliant on formal authority, incentives (e.g. financial), and 'hard' consequences (e.g. termination of employment).	Commits to a high-performance culture by balancing 'hard' performance and accountability mechanisms (targets, performance routines, financial incentives etc.) with 'soft-levers' such as communication, data sharing, and the provision of system support.
<b>Outcome focused targets and supporting indicators</b>	Are not widely understood or owned by the system, used as a punitive measure	Are communicated effectively and are understood by actors at all levels of a system; used to inspire and coalesce a system around a shared outcome
<b>Assessments and quantitative data</b>	Data is not used to identify where support is needed, nor is it shared with the system to inform practice	Data is used to target support and resources where it is needed most, and digestible data is made readily available to the system to inform practice
<b>Targeted and tailored support</b>	Limited support provided to improve system performance, nor is it differentiated, leaving 'strugglers' receiving same level of support as 'high performers'.	Support and resources provided to build system capacity; with specialized and targeted support for struggling performers in the system.
<b>Performance routines and reporting</b>	Over-bureaucratic governance and reporting that is not evidence based nor drives action or decision making.	Frequent data-informed routines and reporting that directly leads to action
<b>System engagement</b>	Site-visits/fieldwork used to evaluate, audit or 'check-up' on frontline staff.	Site-visits/fieldwork used to reinforce priorities, hear from the frontline around what is working well, and identify areas of improvement – with insights to be acted upon.
<b>Public engagement</b>	Inconsistent engagement with public on progress – only when it suits Government.	Trusted and consistent engagement with public on progress – even when performance is off-track.

### 3 SUMMARY AND KEY INSIGHTS

Based on a review of our four chosen case studies, we have identified the following key insights relating to performance management and accountability:

#### System leadership

**Key Insight #1: To drive improvement, leadership requires commitment to a balanced approach to performance and accountability.**

System leaders should recognize that a singular focus on formal authority, incentives, and punitive consequences are largely ineffective in building sustainable system improvement. Rather, any approach needs to balance these with support and engagement that empowers and mobilizes actors across the system to achieve outcome-focused targets.

Central authorities have a role in setting performance standards, yet it is crucial for them to strike a balance between providing guidance and capacity building on meeting these standards and allowing frontline workers the flexibility to respond to the unique conditions they encounter.

Both NSW Education and NSW Health sought to achieve this balance. While outcome focused targets featured in performance agreements of senior leaders and focused performance routines were held regularly with the system leader, these 'hard' accountability mechanisms i.e. formal punitive consequences were not employed punitively but were rather used to encourage problem solving and collaboration. Moreover, frontline engagement (i.e. with teachers and clinicians) alongside the provision of targeted support and resources was also effective in building a strong but supportive performance and accountability culture across these systems.

#### Outcome focused targets and supporting indicators

**Key Insight #2: Targets are most impactful when they are outcome focused.**

Across our case studies, a vision of success articulated in terms of citizen-focused outcomes gave the system direction and focus for its reform efforts. In Punjab, NSW, and Indonesia, we found that education systems track and share learning outcomes as measures of system performance. In Punjab, we found that the system transitioned to using learning outcomes for performance management and accountability once it had got its "basics" right, i.e. addressed challenges of student enrollment, teacher attendance, and infrastructure in schools.

We also found that tying targets to 'high stakes' universal graduation or certification assessments can exacerbate already high levels of stress within the system which can undermine efforts to improve learning outcomes. For example, the NAPLAN test in NSW is a high-stakes assessment which has had concerns raised around its impact on student and teacher well-being. NSW's decision to tie its education target to this assessment resulted in mixed support and acceptance

from the system, particularly from frontline staff. Additionally, it placed further pressure on schools and teachers who were already struggling under the weight of the assessment.

**Key Insight #3: A single outcome measure cannot capture all aspects of system performance; a holistic view is critical.**

Systems cannot rely on a single outcome measure to provide a complete picture of system performance and should consider adopting additional supporting indicators. This provides reference points for validation, ensures that other measures of importance are not being neglected, and also can be used to track unintended consequences. In Punjab, although the monthly Literacy and Numeracy test score was used for performance management, the system continued to invest in more comprehensive six-monthly assessments to understand student achievement. Similarly in NSW Health, while the ‘four-hour rule’ target for patient exit times was the headline indicator for determining hospital performance, several other indicators were also tracked to ensure a holistic view on patient outcomes, including rates of patient readmission and rates of mortality.

**Key Insight #4: A small set of performance measures and targets can enhance focus in a system by communicating the most urgent priorities.**

Performance targets have demonstrated effectiveness when they are used to drive focus and urgency in a system. When too many targets are chosen, they can undermine focus in a system and overwhelm system actors. A small set of targets can effectively communicate the most pressing priorities of a system. In Punjab, for example, a set of 10-12 metrics were used consistently to measure system performance. In contrast, in NSW Health, system actors had over 80 KPIs embedded within service agreements, leading them to continually question where to focus their efforts.

**Key Insight #5: Targets need to be ambitious, but realistic.**

Targets should be differentiated across a system to account for contextual realities – ensuring that targets are ambitious but achievable. This does not necessarily mean holding different parts of a system to different standards, but recognizes that different parts of a system will have different trajectories to achieving the same performance standard.

Several contextual factors outside the control of system actors may affect performance against targets. In education for example, the socioeconomic background of students is a key driver of student learning outcomes. Punjab, for example, controlled for this by differentiating targets for geographical groupings of districts which tended to have similar baselines of student performance i.e. northern districts which historically have more resources and higher household incomes were set a different target from southern and central districts, where the context is different. In NSW Health, hospitals were categorized into four groups based on hospital size and targets were differentiated for each of these.

**Availability and use of data**

## **Key Insight #6: Data underpins a strong performance and accountability culture.**

Data forms the foundation of a strong performance and accountability culture. Not only does data inform targets and their supporting indicators, but it is also used to effectively monitor progress and to focus on where to provide targeted support.

Across our case studies are examples of how data-driven insights can be used to identify localized evidence-based solutions, including understanding what support or resources are needed and to identify which 'best practices' can be replicated. A consistent focus on data that informs action was the key to improvements in Punjab, NSW Education and NSW Health.

### **Targeted and tailored systems of support**

## **Key Insight #7: A strong performance and accountability culture needs an accompanying system of support to drive improvements.**

Effective performance management and accountability requires dedicated support to the system. This includes, for instance, a dedicated function to build system capacity in the use of data to inform practice and targeted support for designing and implementing improvement strategies. In Punjab, for instance, dedicated District Field Coordinators (DFCs) worked with system actors to help them interpret their performance scorecards and devise monthly action plans to improve performance. In NSW Health, the 'Whole of Hospital Team' was established in the Ministry of Health to provide targeted support to underperforming hospitals. The team would embed in struggling hospitals to perform a diagnostic to determine the drivers of declining performance, and then work intensively with the administration and clinicians to drive improvements.

## **Key Insight #8: System equity must be considered when using performance measures to drive improvement.**

As a system focuses efforts on achieving a performance measure or a target, it is critical to ensure that the lowest performers in the system are improving, and that overall improvements are not widening the gap between high and low performers. Reform efforts must consistently target and track low performers to manage this risk.

In NSW Education, for example, the Premier's Priority target aimed to transition average performers into higher performing students. A potential unintended outcome of this target was a deepening chasm between underperforming students, many of whom had a lower socioeconomic background than the rest of the system. In parallel with the Premier's Priority target, NSW also committed to an extensive program aimed at lifting the learning outcomes for the State's most vulnerable schools, as well as tracking the performance of these schools regularly and consistently.

### **Performance routines and reporting**

## **Key Insight #9: Performance routines can accelerate momentum, ownership, and accountability.**

Routines are regularly scheduled and structured conversations between system leaders and accountable leaders to review progress against goals, discuss and solve major challenges, and make decisions to drive implementation. Routines, when executed well, can create momentum and a consistent sense of urgency in the system. Routines are distinct from update meetings and staff meetings: a routine is set up with significant preparation, has a clear 'story' or narrative, is informed by data and has a specific call to action (e.g. decision to be made).

In Punjab and NSW, a key accountability mechanism was regular performance routines, managed by the Special Monitoring Unit and the Premier's Implementation Unit respectively. These routines held system actors answerable to the system leader (e.g. Chief Minister or Premier) for their performance. They also served as platforms to reflect on performance and agree on next steps. In contrast, the Ministry of Education in Indonesia lacks a robust accountability mechanism to hold regional governments accountable due to the decentralized nature of the government. This has meant that despite the availability of performance data, there are limited levers to mobilize action.

### **System engagement and public engagement**

**Key Insight #10: Consistent engagement with system actors can help to embed a performance culture.**

Central engagement with frontline can serve as an effective way to strengthen performance culture across a system. In NSW Education, for example, the Department of Education and the Premier's Implementation Unit jointly conducted regular school visits to make frontline staff 'feel heard', identify barriers and opportunities for achieving the Education target, and to identify and share best practices. The supportive and conciliatory approach ensured that while schools felt a degree of heightened accountability for their performance because of these visits, the visits created a culture of high performance, with the understanding that schools would be recognized for their endeavors and not castigated for any failures.

In addition, engaging the public on progress against a system's goals can create a sense of accountability between the system and those it serves. In the case of the Premier's Priorities in New South Wales (NSW), actively engaging the public and keeping them informed about progress was a crucial component of the performance culture that underpinned the Premier's Priority targets.

## 4 CASE STUDIES

### 4.1 Indonesia (Education) – *System with limited performance and accountability levers*

#### Context and background

Schools: 397,353

Students: 58,063,894

Teachers: 4,095,445<sup>3</sup>

Indonesia's education system is semi-devolved. The National Ministry of Education and Culture is responsible for funding allocation and distribution, national policy, and the design and implementation of national assessments. The delivery of school education is under purview of both Provincial Governments (2<sup>nd</sup> tier of Government) and Regency or City Governments (3<sup>rd</sup> tier of Government).

The Ministry of Home Affairs is responsible for implementing decentralization and administers minimum service standards across a range of sectors including education.

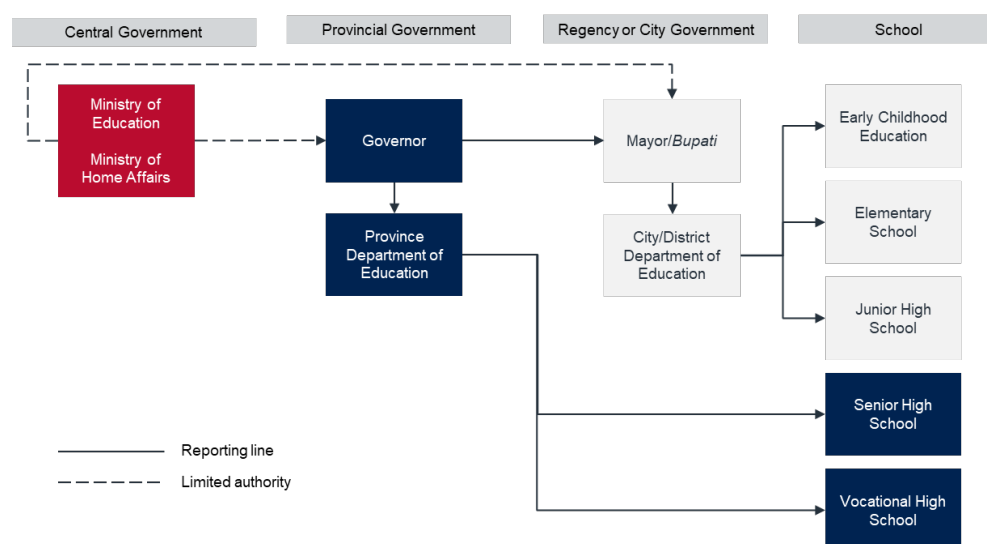
Indonesia has 38 Provinces, 416 Regencies and 98 Cities (administrative areas), making the delivery and administration of education incredibly complex and at times, disjointed.

Provinces oversee Senior Secondary Schools and Vocational Schools, while Regencies or Cities are responsible for Junior High Schools, Elementary Schools, and Early Childhood Education. Each Province, Regency and City has its own local Department of Education.

---

<sup>3</sup> Central Bureau of Statistics (Badan Pusat Statistik) for the academic year of 2022/2023 except for TK for the academic year 2021/2022

**Figure 1. Simplified delivery chain: Indonesia education**



## Target setting and measurement

***After an initial focus on outputs and processes, Minimum Service Standards (MSS) has evolved to include learning outcomes as key indicators.***

The MSS is a set of minimum standards or targets that was initially devised to support Indonesia move towards decentralization. As Provincial and Regency or City Governments transitioned to administering education (and other services), central Government implemented the MSS to hold these sub-governments to account for providing a minimum standard of education to Indonesian children. As the government agency responsible for implementing decentralization, the Ministry of Home Affairs administers the MSS across a range of areas including health, public works and spatial planning, housing, and residential area.

While minimum standard targets for education were initially focused on outputs, covering areas such as the learning environment and access, learning outcome indicators were included in 2021. The inclusion of learning outcomes in the MSS coincided with a renewed focus from Government on the quality of education provided, in addition to coverage and access.

MSS has evolved consistently since its inception in 2001, and while the detail behind the most recent set of standards is largely unavailable, examples of previous (output focused) standards include:

- Each junior secondary school (SMP/MTs) has a natural science lab furnished with desks and chairs for 36 students and a minimum of one set of natural science lab tools for students' demonstration and experiments is available.
- Each primary school (SD/MI) has at least six teachers per school and at least one teacher for every 32 students; and for special regencies, at least four teachers per school.

- Each primary school (SD/MI) has at least two teachers with academic qualifications of a bachelor's degree (S1) or a 4-year diploma (D-IV), and at least two teachers with educator certificates.

The most recent high level MSS indicators are provided below.

**Figure 2. MSS priority indicators for Regency or City governments**

No	Indicators Group	Regency or City Priority Indicators
1	Learning outcomes quality	Literacy competence
		Numeracy competence
2	Learning environment	School safety climate
		Diversity climate
		Inclusivity climate
3	Quality of Early Childhood Education services	Early childhood education services proportion with a minimum accreditation of B
		Proportion of early childhood education teacher with a tertiary education degree
4	Access	School participation rate 5-6
		School participation rate 7-15
		School participation rate 7-18 in equivalency program

***The MSS's design seeks to accommodate for a complex and diverse education system, but the resulting indexed score is difficult to understand.***

Given the sheer scale and diversity of the Indonesian Education System, the MSS is implemented at a sub-government level (Provincial and Regency or City) rather than at a school level. Moreover, the learning outcome standards (literacy and numeracy competence) are designed with a focus on simplicity; stipulating that meeting these standards requires a Province or Regency or City to improve on its National Assessment performance from the previous year.

However, the aggregation of the MSS into an indexed score adds a layer of complexity which makes it difficult to understand, with the policy stipulating its calculation as the '*achievement of indicators of the standard services multiplied by the weights of the indicators*'.

***The National Assessment informs MSS learning outcome indicators.***



The National Assessment informs learning outcome indicator scores in the MSS. Implemented in 2021, the revised National Assessment signifies a shift in educational evaluation in Indonesia, transitioning from a high stakes national assessment assessing individual student achievements to a sample-based evaluation providing a wide-angle view of Indonesia's education system. This shift came about because of concerns of widespread cheating in the high-stakes national assessment. The sample-based approach tests a sample of students in each school, so that representative data on school performance is available. The National Assessment has three components:

- Minimum Competency Assessment (Asesmen Kompetensi Minimum/AKM): Measures literacy and numeracy competencies, focused on assessing critical thinking skills, and is featured in the MSS.
- Character Survey: Measures students' achievements in social-emotional learning outcomes, focusing on character pillars to form the *Pancasila* Student Profile.<sup>4</sup>
- Learning Environment Survey: Is a measurement tool used to evaluate and map the aspects supporting the quality of learning in the schools' environment i.e. factors within the input and learning process.

The Competency Assessment (AKM) is held for a random sample of students in grades 5, 8 and 11 in every school. Unlike the previous National Exam, the AKM results are not to be used to determine student graduation, but as an assessment of school quality.

With exception of their inclusion in the MSS, the Government has mandated no targets for the National Assessment, including no targets for the Competency Assessment (AKM), as these are newly established assessments. The National Assessment results are intended to be used by the school for identifying its strengths and gaps, and as a basis for planning which areas it needs to strengthen. In the absence of clear targets, only highly skilled and dedicated teachers and principals tend to proactively pursue the objective of improving their National Assessment results.

## Performance management tools

***The 'Education Report' is an interactive tool that seek to support performance management at all levels.***

The Ministry of Education shares assessment reports with schools that explain the profile of strengths and areas for improvement for each school and region, through the online *Rapor Pendidikan* (Education Report).<sup>5</sup> *Rapor Pendidikan* is a platform which extracts school-level data from the National Assessment, supplemented by other sources, such as the Central Bureau of Statistics (BPS).

---

<sup>4</sup> Pancasila is the foundational philosophy of the Republic of Indonesia.

<sup>5</sup> Additional information on the Education Report can be found here:

[https://static1.squarespace.com/static/634516bde48db10ce13c8279/t/656063b04738ab5967200954/1700815868252/RaporPendidikan\\_ProductCaseStudy\\_Compressed.pdf](https://static1.squarespace.com/static/634516bde48db10ce13c8279/t/656063b04738ab5967200954/1700815868252/RaporPendidikan_ProductCaseStudy_Compressed.pdf)

The *Rapor Pendidikan* aims to provide schools with the information to identify, reflect on, and ultimately improve learning quality. The platform focuses on school-level performance, showing trends in data over time, and offers recommendations to government officials, school principals, and teachers to strengthen performance against indicators with low scores.

***Results from the National Assessment are not used consistently to support performance.***

National Assessment reports are prepared by the National Ministry of Education capturing changes in achievement at schools over time, based on the sample of students tested at that school. The report does not benchmark performance against other like schools or regions, in recognition of their different resources, conditions, and student cohorts, all factors that influence student learning outcomes. The goal of the Report is to encourage schools to focus on strengthening teaching and learning practices in their schools rather than competing with other schools.

The Report is part of the Ministry of Education's efforts in building the capacity of Provincial and Regency or City governments to understand and use data from the National Assessment. The data can be used, for instance, to understand which school programs are effective and which aren't, and how to strengthen the design of programs. Although the Ministry encourages sub-governments to use the data and support schools, it cannot mandate them to do so. As a result of limited accountability and absence of enforced performance targets, only high-capacity sub-governments and schools make effective use of the data.

**Figure 3. Illustrative of MSS report in the Rapor Pendidikan<sup>6</sup>**

Attainment of priority indicators that affect your SPM index  
Province level table

Priority Indicators	General Senior High School	General Vocational High School	General Junior High School
Literacy Competence	Good	Moderate	Poor
Numeracy Competence	Good	Moderate	Poor
School safety climate	Good	Good	Good
Diversity climate	Good	Good	Moderate
Inclusivity climate	Good	Good	Good
Employment rate of vocational high school graduates		Moderate	

## Incentives and consequences

<sup>6</sup> Source: Adapted and translated from <https://www.govtechedu.id/blog-product-updates/blog-post-rapor-pendidikan-daerah>

***Accountability mechanisms for compliance with the MSS rest entirely with the Ministry of Home Affairs but they are applied inconsistently.***

The MSS is generally framed as 'guidance' for sub-governments in the administration of education and sub-governments are not meaningfully held to account for meeting these standards. However, steps are being taken to strengthen compliance, with the Ministry of Home Affairs starting to tie funding allocation and the approval or endorsement of planning documentation to MSS indicators. Moreover, should the MSS standards not be met, Government regulation stipulates that a Governor (Province) or Mayor (Regency) is liable to administrative sanction, however, this is rarely acted upon as the MSS are not strictly enforced or monitored for compliance.

***The Ministry of Education and Culture has instituted financial incentives for high performing schools, but their impact has yet to be determined.***

The Ministry has implemented financial incentives for high performing schools in the National Assessment, relative to the socio-economic status of their student bodies. The impact of financial incentives is yet to be determined, though there are indications that financial incentives on their own are not effective drivers of performance in the given context. Funds are transferred to eligible schools for a range of uses, including capital investments in schools (maintenance of school facilities e.g. repair of school furniture, repair of computers etc.; investment in school infrastructure e.g. internet, purchase of electricity generators), learning and development or school-based activities, supporting learning and extracurricular activities for students (provision of learning support materials, organizing school student competitions etc.), and funding teachers' professional development activities.

Performance incentives are shared based on four criteria:

- schools with the best National Assessment achievements in each city/district
- schools with the best progress in National Assessment results over the past year in each city/district
- schools with the best National Assessment results from the low socio-economic group
- schools with the best progress in National Assessment from the low socio-economic group<sup>7</sup>

While tying financial incentives to National Assessment results attempts to incentivize performance, there continues to be limited accountability in place around student performance and learning outcomes at a national level. Previously, the National Exam was linked to graduation, making the assessment high stakes for students and their families, teachers and principals. Teachers were held to account by their students and families, teachers were held to account by their principals and so on throughout the school system. It was believed that the pressure placed on the system to achieve results, with limited support, may have led to malpractice and widespread cheating. A study done under the RISE Program in Indonesia provided strong

---

<sup>7</sup> Source: <https://www.kemdikbud.go.id/main/blog/2023/08/lebih-dari-30-ribu-satuan-pendidikan-mendapat-dana-bosp-kinerja-kemajuan-terbaik-tahun-2023>

evidence of wide-spread cheating in the National Exam, including large drops in test scores when “low-integrity” schools transitioned to computer-based testing which prevents cheating.<sup>8</sup> Despite the attempt to limit cheating through the introduction of computer-based testing, scaling up this approach faced obstacles such as computer illiteracy, insufficient infrastructure, and technical issues in test administration.

## Summary of key insights

### *What worked well*

- **Using assessment results to target support:** The current performance management system in Indonesia prioritizes diagnosis, aiming to genuinely understand the prevailing conditions within the education system. By shifting towards a lower stakes environment, the system minimizes incentives for dishonest practices. This intentional design ensures a more accurate representation of the education landscape. The sharing of assessment results aims to engage system actors in genuine improvement efforts without the undue pressure associated with high stakes assessments. This is a step in the right direction, although other aspects of an accountability culture (discussed in the “What didn’t work well” section below) need to be instituted to make this more effective.
- **Adopting a balanced financial incentive mechanism:** The Ministry of Education diversified its incentives to encourage education actors to improve educational quality. This comprehensive range of incentives is intentionally structured to ensure accessibility across all segments of the educational landscape. The different criteria for performance incentives, including separate categories of incentives for most improved schools and schools from low socio- economic groups, ensures that all schools can receive the incentives. By avoiding exclusive distribution to only the top percentile, this approach aims to prevent widening discrepancies and foster a more inclusive distribution of rewards throughout the entire education ecosystem. Moreover, there are specific requirements for how this funding can be spent; to enhance school operations and education delivery. This measure has been implemented to address the potential for malpractice. The full impact of financial incentives is yet to be determined.

### *What didn’t work well*

- **Limited accountability mechanisms in place:** Given the semi-devolved education system in Indonesia, the Ministry of Education lacks accountability mechanisms to hold regional governments accountable. While the Ministry has an incentive scheme in place and makes all National Assessment and MSS data available to regional governments, these have had limited impact in bringing about a cultural shift towards high performance. At present, only regional governments with high commitment and high capacity use the data and recommended actions available through the Education Report to undertake school improvements.

---

<sup>8</sup> Source: <https://riseprogramme.org/blog/computers-indonesia-cheating-learning>

- **Undifferentiated targets:** Results of the National Assessment are shared with sub-governments to help them understand school performance. The communicated goal is for schools to improve their results compared to the previous year. However there are no clearly defined targets for sub-governments or for schools. One of the reasons for this is that the National Assessment is relatively new, and there is an ongoing effort to establish a baseline. There is no indication by the government at present on setting targets in the future, In the absence of clear targets and an accountability culture, only highly skilled and dedicated teachers and principals tend to proactively pursue the objective of improving their National Assessment results.

## 4.2 New South Wales (NSW) (Education) - System with mixed performance and accountability levers

### Context and background

Schools: 2200

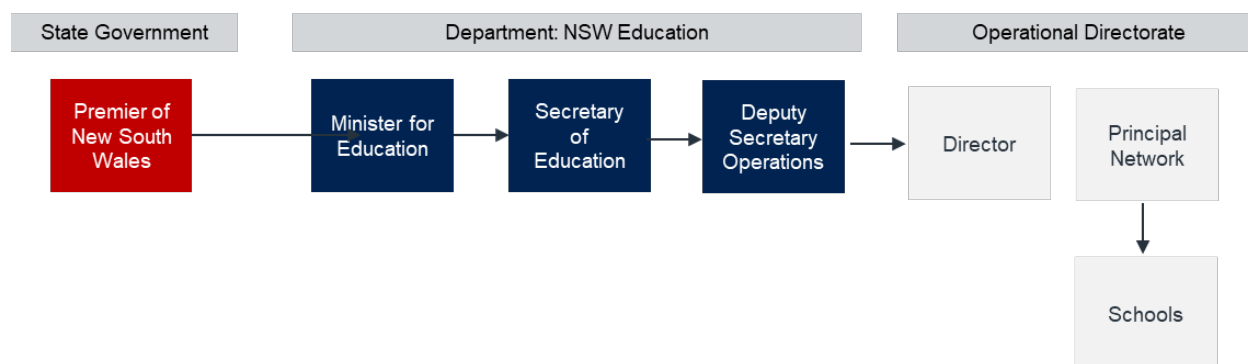
Students: 1.2 million

Teachers: 94,000<sup>9</sup>

New South Wales (NSW) public education system is the largest in Australia, encompassing 2,200 schools and 800,000 students. It is headed by the NSW Department of Education as the central agency overseeing eight Regional Operational Directorates. NSW Education is largely centralized, with the Department of Education responsible for the overall management of the public school system. The Operational Directorates are responsible for supporting schools in their jurisdiction.

Within directorates, groups of schools are grouped into “Principal Networks”, headed by a Director. These networks are a part of the professional development and support structure within the NSW education system, and the Director acts as a coach for the school principals in his network.

**Figure 4. Simplified delivery chain: NSW education**



This case study will look at performance and accountability practices in NSW Education with relation to the Premier’s Priorities (2015-2019), a set of time-bound, citizen-oriented targets that were introduced across various key sectors by the head of NSW state – the Premier. To support the delivery of priorities, the Premier’s Implementation Unit was set up in 2016. The Unit worked closely with frontline teams on data-led, evidence-based interventions to improve results, and supported monitoring of progress against targets.

<sup>9</sup> Source: <https://education.nsw.gov.au/>

## Target setting and measurement

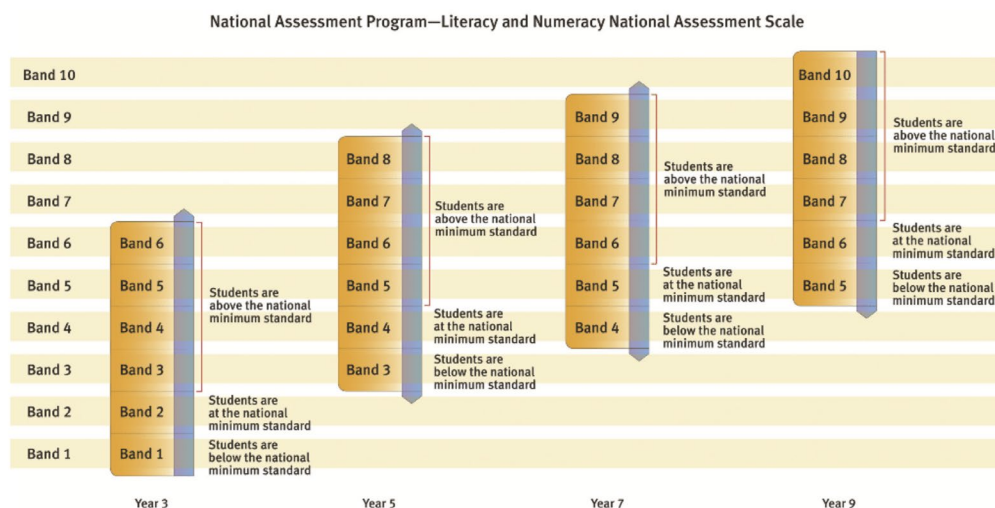
***The Department of Education set an outcome focused target for improved learning outcomes as part of the Premier's Priorities.***

A key measure of student learning outcomes across Australia is the National Assessment Program – Literacy and Numeracy (NAPLAN) test. The NAPLAN is an annual universal assessment of students in Years 3, 5, 7 and 9 across five areas: reading, writing, grammar and punctuation, spelling and numeracy.

NAPLAN results are divided into ten bands, with each band representing a different level of student proficiency. Band 1 is the lowest band and band 10 is the highest band, and there is a minimum band defined for each year of schooling (Figure 5). The band range—from band 1 to band 10—reflects the increasing complexity of skills and understandings demonstrated by a student and assessed by NAPLAN testing as the student progresses from Year 3 to Year 9. For any one year level, the full range of student performance is reported using six of the ten bands. For example, student performance at Year 3 is reported within the range of bands 1 to 6, whereas Year 9 student performance is reported within the range of bands 5 to 10.<sup>10</sup> This type of assessment that categorizes students into a series of bands that stretch across grades is particularly valuable for monitoring system performance across grades and time, as opposed to distinct assessments for different grades which cannot be used for this type of system tracking.

In 2015, the NSW government announced 12 Premier's Priorities for the period 2015-19, across a range of key areas important to citizens, including education. The target for education was to “increase the proportion of NSW students in the top two NAPLAN bands by 8% by 2019”.

**Figure 5. NAPLAN assessment scale<sup>11</sup>**



<sup>10</sup> Source: [https://docs.acara.edu.au/resources/Fact\\_Sheet\\_-\\_About\\_NAPLAN.pdf](https://docs.acara.edu.au/resources/Fact_Sheet_-_About_NAPLAN.pdf)

<sup>11</sup> Source: [https://docs.acara.edu.au/resources/Fact\\_Sheet\\_-\\_About\\_NAPLAN.pdf](https://docs.acara.edu.au/resources/Fact_Sheet_-_About_NAPLAN.pdf)

### ***The decline of NAPLAN's evolving reputation significantly impacted support for the target.***

NAPLAN was introduced in 2008 and its purpose was to “monitor progress towards national goals, school system accountability and performance, school improvement, individual student learning achievement and growth, and information for parents/carers on school and student performance”.<sup>12</sup>

Unveiled with much fanfare, NAPLAN was viewed as the panacea to education in Australia, which was largely seen to be disconnected and with a focus on the short term. In the early years NAPLAN was seen to be fulfilling many of its stated objectives. It provided a standardized and rigorous approach to assessment, facilitating a national view on student achievement and supporting Government to take an informed and consistent approach to school improvement.

However, over time, a range of unintended consequences began to emerge, stemming from the 2010 decision to publish NAPLAN results to the public. These results were widely used by the public, particularly the media, to make judgements of systems, schools, and teachers, and to compare non-government and government education. As a result, the NAPLAN assessment became high stakes in nature leading to detrimental effects on student wellbeing and on school staff morale. Concerns about the effects of NAPLAN on student wellbeing were raised by parents and educators, with students experiencing high stress levels due to NAPLAN tests. There were also concerns about narrowing of teaching strategies and of the curriculum, negative impacts on staff morale, negative impacts on school reputation and capacity to attract and retain students and staff.<sup>13</sup> While the NAPLAN assessment was seen to be a rigorous tool in determining learning outcomes, its reputation suffered due to these concerns. .

System buy-in and support for the NAPLAN target was impacted by the assessment's reputation. While teachers and principals recognized the benefit of having a singular focus on learning outcomes to prioritize investment, teaching practice, and support, tying it to the NAPLAN assessment weakened support.

### ***The target was difficult to understand which limited its impact.***

When used well, targets can be effective at galvanizing a set of actors around a common goal. This can only happen with a clearly understandable target, one that can be communicated clearly throughout a system, from senior decision makers to frontline staff, to the community. Many were unable to easily understand the NAPLAN target, with even the Minister consistently requesting clarifications. Examples of the points of confusion include (those who most requested clarification in brackets):

- ‘Top two bands’: unclear on what is meant by a band (community)
- ‘8% increase’: confusion around an increase from what to what (Minister, teachers and community)

---

<sup>12</sup> Source: [https://naplanreview.com.au/pdfs/2020\\_NAPLAN\\_review\\_final\\_report.pdf](https://naplanreview.com.au/pdfs/2020_NAPLAN_review_final_report.pdf)

<sup>13</sup> Source: <https://lens.monash.edu/@education/2021/05/10/1383196/learning-from-disruption-why-we-should-rethink-the-place-of-naplan-in-our-schools>



- 'NSW Students': confusion about which grades were included (Minister, teachers and community)

Finally, the baseline for the target was particularly complex, comprising of the average of 2014 and 2015, rather than using a single year.

The lack of clarity around the target made it incredibly challenging to communicate it through the system, to help schools understand how they were doing against the target, and to tailor and provide targeted support (including analysis) to schools and principals. These problems underscore that effective communication of targets is as important as setting a technically sound target.

***System equity was considered when setting the target.***

The target's intended outcome was to transition average performers into higher performing students. It was expected that most, if not all, of the 8% growth in the top two NAPLAN bands (the stated target) would come from students in the middle NAPLAN bands. An unintended possible outcome of this target was a deepening chasm between underperforming students, many of whom had a lower socioeconomic background than the rest of the system, since there was no explicit target for improved performance among students in lower NAPLAN bands.

As a result, NSW Education committed to lifting the learning outcomes of vulnerable schools and students. The performance of these schools and students was regularly monitored by reviewing student results to ensure their rate of improvement matched that of the broader student population.

One way to control for equity considerations is to include explicit equity targets in a reform effort, as in the case of the No Child Left Behind (NCLB) in the U.S.<sup>14</sup> This ensures that neglected segments of the population are a constant focus in the reform effort, and are receiving targeted support and resources. Often, the system can see overall gains in its performance by directing its resources to these segments.

## **Performance Management Tools**

***Rather than using the NAPLAN target as a punitive measure, the Department focused on providing support informed by data.***

The NSW Department of Education took a targeted and data-driven approach – “Bump It Up” – to lifting results, where for the first time ever, data was used to identify 137 schools with a large proportion of students in the middle NAPLAN bands, that were underperforming compared to

---

<sup>14</sup> An example of a reform that embedded focus on improving performance of the lowest performers is the “No Child Left Behind” reform in the US, which explicitly required that progress be tracked and strengthened for groups disaggregated by socioeconomic status and by ethnicity.

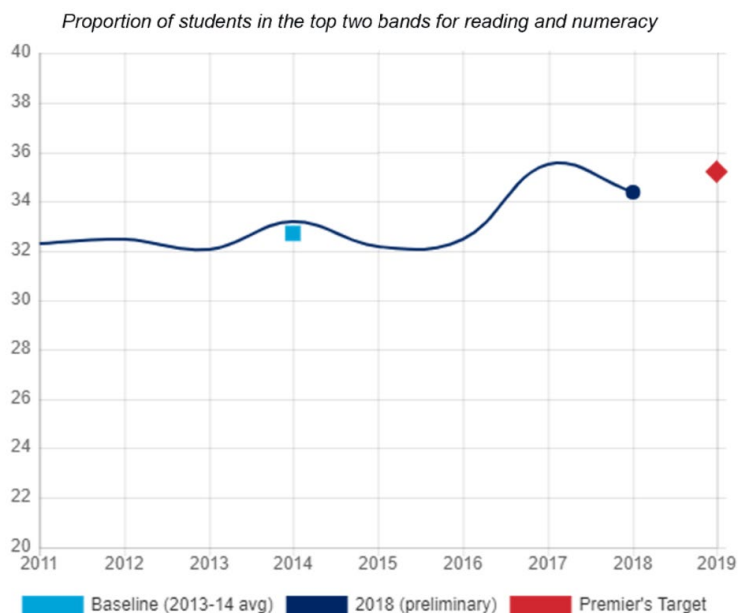
statistically similar schools (based on remoteness, level of socio-educational advantage and other factors).

Under the Bump It Up Strategy, these schools were provided focused support to improve NAPLAN scores. This included support to use individual school data to develop targeted reading and numeracy initiatives. Principals were tasked with developing a plan to lift their literacy and numeracy results based on their school-level data and using evidence-based practices. The Department of Education and the Premier's Implementation Unit supported principals and teachers in the development and implementation of these plans, as well as identifying and sharing good practices across the 137 schools for their adoption. No financial support was given to Bump It Up schools.

The Bump It Up Strategy led to Bump It Up schools improving literacy and numeracy performance at twice the rate of non- Bump It Up schools. The strategy contributed to the Premier's literacy and numeracy target being achieved two years ahead of schedule in 2017. A subsequent shift in policy of phasing out the Bump It Up strategy in favor of a universal support model where all schools were provided the same access to support, as well as leadership change at the highest levels of Government, may have contributed to a decline in performance in 2018.

Overall, the proportion of students in the top two NAPLAN bands for reading and numeracy is estimated to have increased from ~32% in 2011 to ~36% in 2017.

**Figure 6. Bump It Up strategy and policy shift effect**



## Incentives and consequences

### ***Regular routines held senior Education officials to account.***

Regular data-informed routines focused on action and decision making were held between the Premier, Minister for Education and senior Department officials. Every month a short and engaging update note was developed for the Premier and Minister that outlined past and projected performance against the education target, as well as next steps for the following month. This Note was written in partnership between the Department of Education, which provided education expertise, and the Premier's Implementation Unit, which provided delivery expertise and an independent viewpoint.

Quarterly Stocktake meetings were also held between the Premier, Minister and Education Officials. These were 1.5-hour long meetings held in person, dedicated to genuine deliberation around progress and problem solving, with a focus on unlocking barriers to delivery via engaging decision-makers in the room. The Stocktake meeting required meaningful preparation from both the Department of Education and the Premier's Implementation Unit to ensure decisions were informed by data and evidence of progress on the ground.

### ***'Fieldwork' was conducted regularly to engage the system and identify best practice.***

Both the Department of Education and the Premier's Implementation Unit jointly conducted regular school visits to reinforce the Premier's priority of improving student learning outcomes. The visits were intended to engage the system and make frontline staff 'feel heard', to identify barriers to and opportunities for achieving the Education target, and to identify and share best practice across the system.

The visits were not an exercise of 'checking up' on schools or passing on blame for poor performance, but rather an opportunity to meaningfully engage the system on what's working well and what can be improved. These visits created a culture of high performance, where schools became invested in trialing innovative interventions, with the understanding that schools would be recognized for their endeavor, and not castigated for any failures.

Fieldwork visits were often made with limited notice to schools, and in a handful of cases, the Minister also joined. Fieldwork was always conducted in a way that limited disruption, with different schools selected for each fieldwork cycle. Schools were selected based on a set of hypotheses developed prior by the Department of Education and the Premier's Implementation Unit. These hypotheses centered around performance and took into account high performers, low performers, socio-economic status, indigenous student cohort and location.

The supportive and conciliatory approach to fieldwork ensured that while schools felt a degree of heightened accountability for their performance because of these visits (or the prospect of a visit), they were generally welcomed by principals and teachers.

## **Summary of key insights**

### ***What worked well***

- **Targeted support for schools which needed it the most:** A key contributing factor to the success of the Bump It Up strategy was its focused support to schools which needed it the most and were underperforming compared to statistically similar schools. This ensured that the system was directing resources where they could add the most value.
- **Evidence-based strategies to improve school performance:** Strategies for improvement of test scores were developed locally by schools, grounded in an understanding of their school-level data. To support schools, the system also shared information on evidence-based best practices most aligned to the challenges faced by schools. This not only resulted in strategies being effective, but also built the capacity of schools to understand and address their challenges independently.
- **Regular ‘fieldwork’ conducted to engage the system:** Fieldwork visits conducted jointly by the Department of Education and the Premier’s Implementation Unit played a key role in engaging the education system, as well as identifying and sharing best practices that directly contributed to improved learning outcomes.

#### ***What didn’t work well***

- **Setting a target that was difficult to understand:** The NAPLAN target was difficult to understand and consequently difficult to communicate effectively to the system. In addition to the baseline being unclear, it was also not clear how many children the “8%” represented. This made it challenging to coalesce the system around the learning target and particularly challenging to engage the public. As a result engaging the system around improvement strategies at times lacked focus and direction.
- **A singular focus on NAPLAN to inform the target:** Setting a sole learning outcome target from NAPLAN assessment results may have exacerbated pre-existing issues that were anecdotally linked to the NAPLAN assessment. Student wellbeing, narrowing teaching practice and curriculum, and increased teacher burden were all emerging issues related to the high public visibility of the NAPLAN. Tying a public, system-wide target to this assessment likely embedded these issues.

### 4.3 Punjab, Pakistan (Education) - System with strong performance and accountability levers

#### Context and background

Schools: 52,000

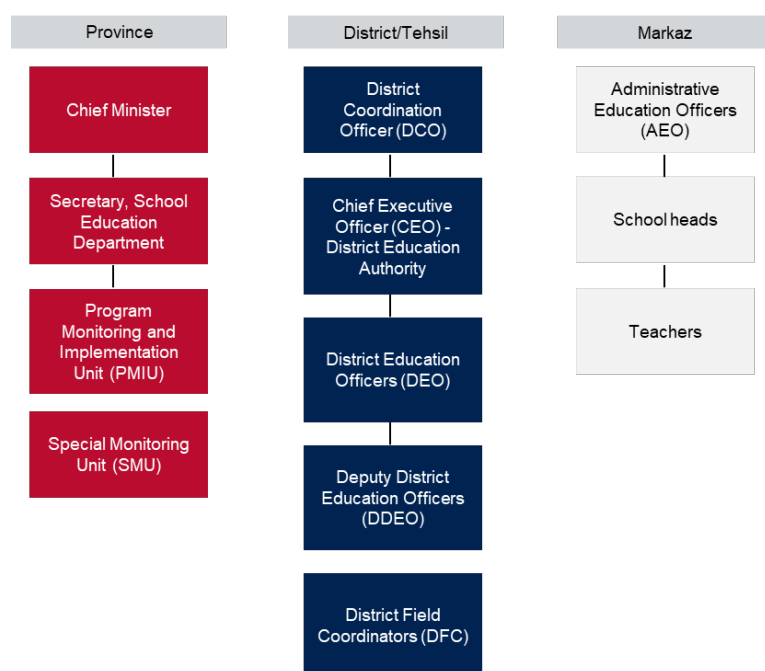
Students: 12,000,000

Teachers: 400,000<sup>15</sup>

Punjab is the most populous province in Pakistan, with a population of a 110 million people. Education is devolved in Pakistan, meaning that Punjab functions autonomously of the federal government. Punjab is divided into 36 districts that are further divided into *tehsils* and *markaz*. Each district is headed by a District Coordination Officer (DCO) and has a District Education Authority (DEA) headed by a Chief Executive Officer (CEO). District Education Officers (DEO) report to the CEO, and Administrative Education Officers (AEO) report to DEOs.

A simplified overview of this delivery chain is illustrated in Figure 7.

**Figure 7. Simplified delivery chain: Punjab education**



<sup>15</sup> Source: Punjab School Census 2017

Punjab introduced a comprehensive performance management and accountability system as part of the School Reforms Roadmap (2011-2018), an ambitious reform agenda which was personally overseen by the Chief Minister of the province. The Roadmap had three core priority areas: increasing enrollment, improving education quality and learning outcomes, and providing physical and human resources to schools. Across these three areas, the Roadmap rolled out multi-pronged strategies to address challenges faced by the system. System-wide performance management and accountability was underpinned by the establishment of outcome targets, regular collection and sharing of data, data-based performance monitoring, and support for underperforming regions and schools.

These reforms achieved success over the several years they were implemented, including an increase in the Participation Rate (percentage of 5-9 year olds that go to school) by 5.7 percentage points between 2011 and 2017, from 84.8% to 90.5%, which translated to nearly an additional 1 million children in school.<sup>16,17</sup> Between 2015 and 2017, learning outcome scores based on the six-monthly assessment, which tested Grade 3 students on a set of Student Learning Outcomes (SLOs) across English, Mathematics, and Urdu, showed an increase from 56% to 68% of percentage of correct answers on the test, which was a substantial gain.

## **Target setting and measurement**

### ***Targets were set top-down, with limited consultation with Districts.***

Targets for all indicators were set by the Roadmap team in consultation with officials from the Schools Education Department, based on baseline data and projections of improvement (“trajectories”). District CEOs were not involved in determining the targets, a decision taken to prioritize action over the slow process of “buy-in”. Once targets were set, all CEOs were invited to the provincial capital where targets were shared by the Secretary. CEOs were responsible for sharing the targets with the education actors in their districts, who communicated them to schools.

Over the years of the Roadmap, targets were reevaluated in consultation with provincial education officials to ensure that they continued to be both realistic and ambitious.

Because the targets were perceived as top-down, over the years the system struggled with instilling a sense of ownership of targets among system actors. This meant that system actors often referred to the targets as unrealistic and unachievable, although actors at the central government viewed the targets as achievable

### ***Significant investment was made in data collection, analysis and engagement.***

Data to measure performance against targets was collected monthly for a large sample of schools in all districts by independent data collection officers. The Program Monitoring and

---

<sup>16</sup> The Participation Rate measures the number of 5 to 9 year olds in school as a percentage of the total population of 5 to 9 year olds in Punjab. This was chosen as the headline measure of access in Punjab, as a simpler and more contextually relevant measure than Net Enrollment Rate (NER) or Gross Enrollment Rate (GER) given the early stage of educational development in Punjab.

<sup>17</sup> Based on data from the Punjab School Education Survey (PSES) conducted by Nielsen

Implementation Unit (PMIU) was a dedicated unit at the School Education Department which was responsible for collecting and analyzing the data. PMIU employed nearly 1100 Monitoring and Evaluation Assistants (MEA), who were assigned to collect data from 40-50 schools each month (this collectively constituted a very large random sample of schools in Punjab). MEAs would make unannounced visits to schools and record data on a tablet-based application, which uploaded the data directly to a central server. The set-up of this data collection system including human resource and technology was the result of a massive financial investment by the Punjab government and its partners. The team at PMIU would analyze this data to develop “scorecards” and share these with the district officials.<sup>18</sup>

***The Roadmap prioritized a set of 10-12 indicators that were consistently tracked for improvement.***

These included indicators on school access (student attendance, student retention, teacher presence), school infrastructure (functioning facilities, presence of boundary walls, sufficiency of toilets) and student learning (LND scores). There were detailed standard operating procedures (SOPs) in place on how this data was to be observed/collected by MEAs during their school visits, as well as standardized formulas for how the indicator value was calculated. For example “availability of boundary wall” was marked “yes” if a school had a boundary wall that was 5 ft high, surrounded the school completely, and had no gaps.

While keeping focus on a set of indicators consistently was good, district officials would express concern that external factors that were not considered in the analysis sometimes prevented them from showing improvement against indicators. For example, sometimes recruitment for open teaching positions would be slow, or funds requested to fix school infrastructure would be delayed.

***Learning outcome targets were only set once Punjab got the ‘basics’ right.***

As part of the Roadmap reform strategy, targets for key indicators were introduced for all districts of Punjab across the core priority areas of the Roadmap, including for student enrollment, school facilities, and learning outcomes. Targets for learning outcomes were introduced a few years after the launch of the Roadmap, as in the earlier years the “basics” were prioritized i.e. student enrollment, teacher attendance, and functioning facilities at schools.

At the outset it was recognized that the Punjab education system was in the early stages of maturity, and that a focus on learning outcomes would be largely ineffectual without teachers to teach, classrooms to teach in and students to learn. After the first three years of the reform, with many of these foundational elements mostly in place in most of Punjab’s districts, the system shifted towards a learning outcome focus.

At the time there was limited evidence on learning outcomes in Punjab. The province administered an annual grade 5 exam, but there was no comprehensive measure of learning in the early grades. In 2014, the system conducted the “Six-monthly assessment” in a sample of schools in

---

<sup>18</sup> Punjab’s set up of its data collection system is detailed in its publication “*Transforming primary education in Punjab: The journey of Punjab’s Education Reform Roadmap*”.

the country, which found that students were not learning commensurate to their grade level. The province introduced an integrated approach of reforming curriculum, textbooks, teaching support, and assessments to improve learning outcomes. To measure the ongoing impact of this reform, the following year the government piloted the Literacy and Numeracy (LND) test as a monthly measure of school learning.

The LND test was administered by the MEAs during their school visits. MEAs would select a random sample of 6 students in every Grade 3 classroom visited as part of their assigned monthly school visits, to take a tablet-based test based on a set of language and mathematics concepts from the Grade 1 and 2 curricula. This allowed for a monthly measure of learning outcomes. The LND test measured specific student learning outcomes (SLOs) for English, Math, and Urdu. LND scores were reported as a percentage of correct answers on the LND test. LND scores were the only indicators against which scores were not made available at school-level since the measure was not statistically significant for individual schools.

The LND was chosen as the learning outcome metric that would be monitored and tracked as part of the Roadmap. While there were a range of learning improvement areas that required urgent attention, and learning reforms were focused across primary grades, the Government selected the LND metric for Grade 3 learning outcomes in order to drive focus and to not overwhelm the system.

***Targets were differentiated across districts for some indicators, to ensure that they were ambitious but achievable.***

For some indicators, such as classroom observations and school facilities, targets were set at the provincial level and uniformly applied to all districts. However, for some of the other indicators such as student attendance and LND, targets were set for geographical groupings of districts which tended to have similar baselines i.e. for northern, central, and southern Punjab.

Effective targets typically balance pragmatism with ambition. District level targets sought to achieve this by setting an aspirational benchmark while accounting for contextual factors such as historical performance, demographics, and the level of resourcing in place.

## **Performance management tools**

***The ‘Performance Scorecard’ was the primary tool used to manage performance.***

The primary performance management tool was a “performance scorecard” generated monthly by the PMIU for each of the 36 districts in Punjab, based on the data collected by MEAs during their school visits. The performance scorecard comprised the key indicators aligned with the Roadmap priorities against which targets were set, including school attendance, learning outcomes, and school infrastructure. Scorecards were disaggregated at different administrative levels, including district, *tehsil*, and *markaz* levels. This allowed the scorecard to be used to capture the “performance” of district CEOs, DEOs, and AEOs. For example, the *markaz*-level scorecard for the AEO would show indicators for all the schools assigned to an AEO (except LND measures, as stated above). Similarly, the *tehsil*-level scorecard would show indicators for a



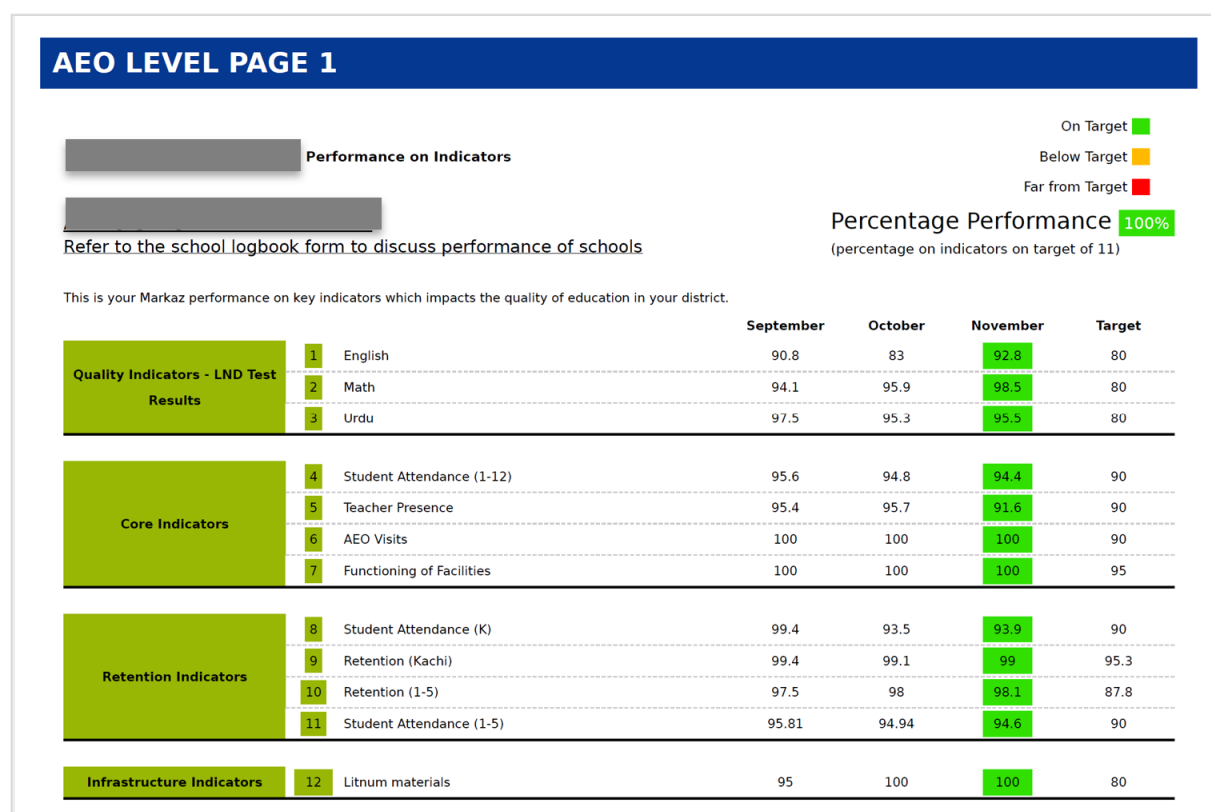
group of *markaz* i.e. *tehsils* assigned to a DEO. Illustrative performance scorecards for AEOs are shown in Figure 8.

***Targets were not explicitly used to performance manage teachers or principals.***

At the school level, performance management of teachers and principals did not explicitly draw on targets. The rationale for this approach was twofold. Firstly, it was believed that the use of targets was not part of the cultural norm for school staff, and that that their rigid application would lead to significant resistance and broader discontent from the workforce. Secondly, and more critically, explicitly tying individual teacher performance to learning outcomes was thought to be highly questionable given the range of exogenous factors that directly impact learning outcomes (e.g., demographics and socio-economic status, home-life and healthcare). These factors are particularly influential at an individual, classroom or school level, but can be largely accounted for at a system level through differentiated targets and the tracking of additional indicators that provide a more holistic picture of performance.

While AEOs were responsible for supporting teachers and principals with teaching and school management skills with the aim of improving their scores on Roadmap indicators, there was no formal performance management to hold teachers or principals to account for Roadmap targets,.

**Figure 8. Illustrative performance scorecards for AEOs**



## AEO LEVEL PAGE 2

School wise performance

On Target ■  
Below Target ■  
Far from Target ■

Refer to the school logbook form to discuss performance of schools

This is a school level performance of your Markaz on various indicators.

Indicator #	1	2	3	4	4A	5	6	7	8		
School	Student Att. (1-12)	Teacher Presence	AEO Visits	Functioning of Facilities	Partially Func. Facility	Student Att. (K)	Retention (Kachi)	Retention (1-5)	Student Att. (1-5)	Performance in current month	Past 3 Months Average
	96.9	90.9	100	100	-	92.3	100	99	95.2	100	95.8
	100	100	100	100	-	100	100	100	100	100	100
	90.6	100	100	100	-	90.4	100	98	90.6	100	100**
	93.6	100	100	100	-	100	100	98	93.6	100	100
	95.2	100	100	100	-	100	100	100	95.2	100	100
	96.2	100	100	100	-	100	100	100	96.2	100	100
	94	50	100	100	-	95.9	100	97	95.3	87.5	87.5

## Incentives and consequences

Within a district, all district officials, including the DEOs, DDEOs, and AEOs were held accountable against the district targets. Although targets were applicable to all schools, school officials were not individually held accountable to targets and did not participate in accountability routines (as detailed below). AEOs were responsible for communicating and discussing performance with schools and supporting them to improve performance. Punjab does not have a legislated system for performance management of school officials.

**High stakes monthly and quarterly routines were established to hold administrators to account.**

The primary accountability mechanism for performance scorecards was a cadence of “routines” – structured meetings with a focused agenda on reviewing indicators, holding district officials accountable for their performance, and problem-solving on low performance. A district-based performance management cycle was instituted monthly, gathering actors to review the month’s performance scorecards and to follow up on actions agreed in the previous month’s meetings. These routines allowed the head of the district – the DCO, and the Education CEO to hold all actors reporting to them accountable for their performance against the indicators. District routines

created accountability by establishing a platform for sharing of data among all the participating district actors, with their peers in attendance.

Every quarter, data on all districts was presented to the Chief Minister in a Stocktake meeting, where education officials had to answer for their district's performance against targets. The Stocktake was a high stakes meeting attended by key provincial ministers. The Chief Minister would personally take note of high and low performing districts, and poor performance could result in demotion of district heads. The Stocktake created a strong incentive for district managers to show good performance against set targets, and this translated to pressure to meet targets throughout the delivery chain.

A deep commitment from leadership (both the Chief Minister and District CEOs) underpinned the effectiveness of routines, ensuring they were never rescheduled and that attendees were sufficiently prepared to facilitate informed deliberation and decision making.

***Financial bonuses used to incentivize performance led to perverse outcomes.***

The monthly performance of each district was also used to produce a quarterly district ranking of all districts. The DCO and CEOs of the five top performing districts would receive a generous performance bonus which created a strong monetary incentive for them to ensure district officials reporting to them were meeting performance targets. No monetary incentives were offered to other district officials or to school heads/teachers. This resulted in DCOs and CEOs exerting undue pressure on district actors who reported to them to improve performance across Roadmap indicators.

The district ranking and high stakes Stocktake meetings led to perverse incentives in the system. For instance, to improve LND scores, schools would actively teach to the small set of student learning outcomes (SLOs) included in the LND test, schools would assign the best teachers to grade 3 for which the LND test was conducted, and schools would sometimes even resist enrolling weak students in grade 3 to maintain their LND scores.

***Strong accountability was balanced with support to Districts to improve performance.***

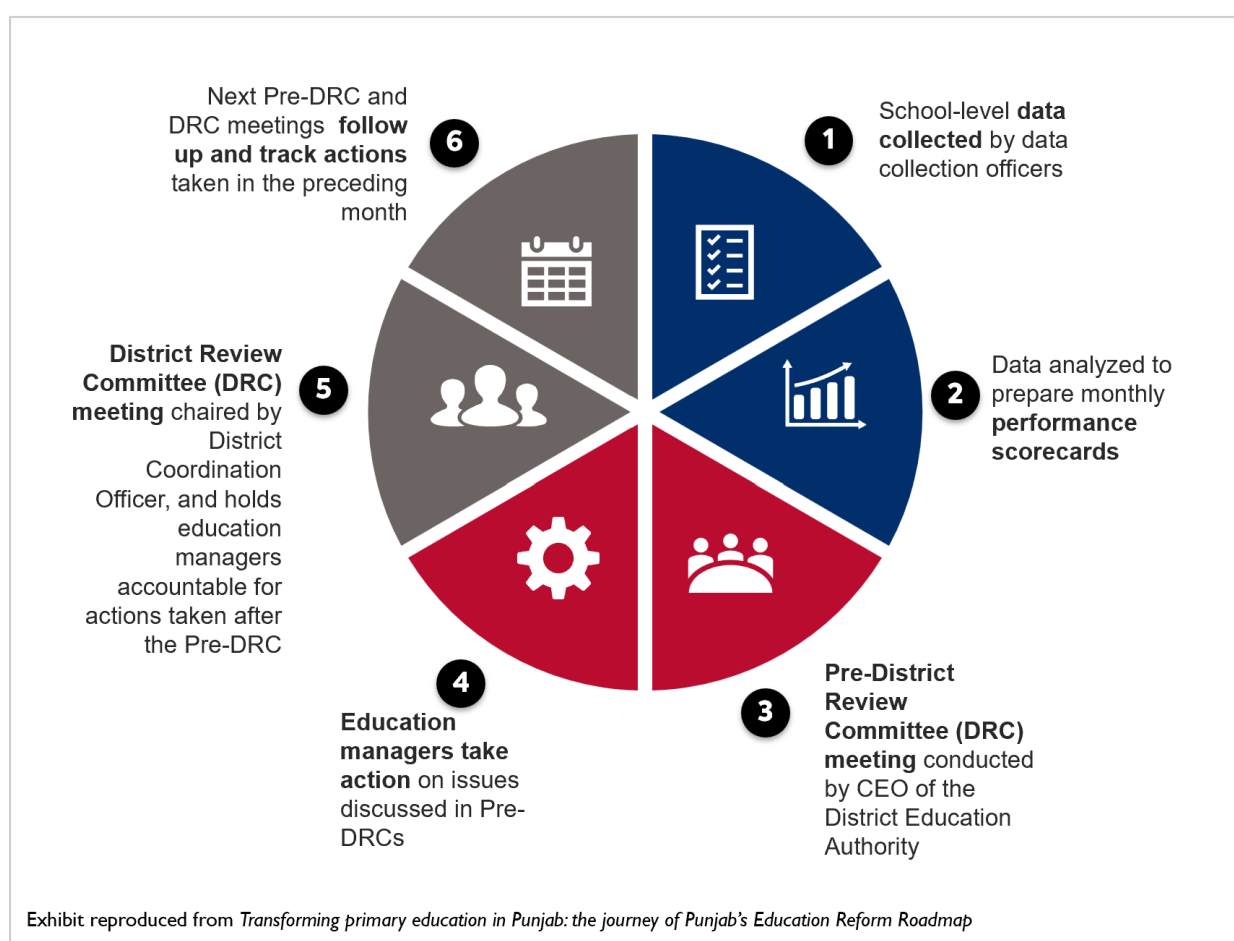
In addition to establishing accountability, district routines also offered the opportunity to solicit targeted support to improve performance. Following these meetings, district officials would develop workplans to improve performance in weak areas. In addition, CEO's would have dedicated meetings with the 4 weakest performing AEOs in their district, in which AEO's would present a plan of SMART actions to improve performance of their *markaz*.

District Field Coordinators (DFCs) were assigned across districts to support the management of routines. Each DFC was assigned to support 4 districts. Apart from managing the routines, DFCs played a key role in supporting district officials to improve their performance across the performance indicators. This included understanding the root causes for problems, identifying solutions together with district actors, and escalating problems to upper management as needed. The role of the DFCs ensured that the performance management structure in the system provided

both challenge and support. As a follow-up to the district meetings, for instance, DFCs were responsible for supporting AEOs with drafting improvement plans for the following month.

As part of targeted support, the Roadmap also strengthened the “middle tier” of district management by improving the capacity of Administrative Education Officers (AEOs) who were responsible for providing support to school heads and teachers. This involved increasing the number of AEOs so that each AEO was responsible for a manageable number of schools, strengthening AEO training, and instituting dedicated DFC support for AEOs. AEO support to schools included conducting classroom observations and identifying professional development needs for teachers.

**Figure 9. Monthly District performance management cycle**



## Summary of key insights

### What worked well

- **Implementing effective performance management routines:** Regular performance management routines based on targets created a strong accountability culture in the province, driven by the Chief Minister, and involving key system administrators such as

AEOs. This accountability structure driven by sharing of data, district-based routines, and the involvement of system leaders such as the Chief Ministry and Secretary of Education created momentum for change and encouraged complacent district leaders to understand the issues at school level and to solve these issues. The accountability structure proved to be a strong catalyst for behavior change, for e.g. AEOs who had been neglectful became vigilant with their school visits.

- **Investing in data systems:** The availability of regular, reliable, and granular data was the primary enabler that underpinned the performance management and accountability cycle. Due to the investment in an independent data collection system, data collected every month was immediately available on a central database for analysis. Monthly performance scorecards were made available to district education actors at all levels, with easy-to-understand analyses. This monthly cycle of data collection and review aimed to establish a sense of urgency and momentum around the Roadmap reforms, by bringing system actors together frequently to discuss the most up to date data. Most importantly, the availability of data at the school level meant that it could be used for accountability of actors at different levels, aggregated and disaggregated as needed.
- **Ensuring dedicated support for district actors:** The performance management and accountability measures were accompanied by dedicated support to district officials by DFCs, and dedicated support for schools by AEOs. This system of support ensured that district officials had guidance, training, and help in understanding how they could strengthen performance against indicators. The strengthening of the AEO role meant that schools were given the support they needed to be able to drive improvements in performance. These systems of support were enablers for performance targets to positively impact system performance.
- **Ensuring dedicated capacity to manage performance routines:** Performance management routines were made possible by dedicated capacity in the system for managing the routines. Significant preparation would go into each routine to ensure that meetings were focused on the most pressing issues, participants had the chance to understand and review their scorecards in advance, and participant were prepared to talk through their challenges and the help needed from the system. DFCs were assigned across districts to support the management of routines. As these routines were implemented throughout the system, there was variability in their application – with some regions implementing them with rigor while others had limited capacity and capability to do so effectively.

### ***What didn't work well***

- **The emergence of perverse incentives:** Although the performance management system led to improvements in school performance over time, the use of high-stakes targets linked to monetary incentives, demotion, and chastisement by the Chief Minister himself led to a number of perverse incentives among schools and district officials. For the indicators on learning outcomes, for example, there were reports of “teaching to the test”, where

teachers would dedicate a large amount of classroom time to the learning outcomes that would be tested in the monthly data collection visit. Schools would also try to influence the student sample that took the tablet based LND test, would assign their best performing teachers to grade 3 to be able to improve LND scores, and would try not to admit weak students to grade 3 (putting them in grade 2 instead). These continued to be problems in the system despite regular data audits and triangulation of learning outcomes data through other system assessments.

- **Instituting top-down targets:** “Top-down” targets were sometimes cited as unfair by district officials, especially given the differentiated targets among districts for certain indicators. Sometimes district officials claimed that they were being held to account for indicators which they did not have power to influence. For example, for school infrastructure needs, their requests for funds were sometimes pending for several months, leading them to show consistently poor results in those areas despite having taken the action they could.
- **Routine fatigue:** Over time, the district routines lost some of their initial value-add. The quality and effectiveness of routines was dependent on the capacity of the DFC and how effectively he/she was able to engage the district leaders. Consequently the usefulness of the routines varied across districts. The high frequency of routines (monthly) meant that sometimes the same issues would be brought up in each monthly meeting, as one month was too short a time to show substantial improvements or to have taken action on tasks from the previous month’s meeting. Over time as the pace of reform slowed down, district actors began to express fatigue with the high frequency of routines, and with the stress that came from being held answerable for low performance in front of their peers.

#### 4.4 New South Wales (Health) - System with mixed performance and accountability levers

##### Context and Background

Public hospitals: 228

Staff: 131,866

Annual presentations: +3,000,000

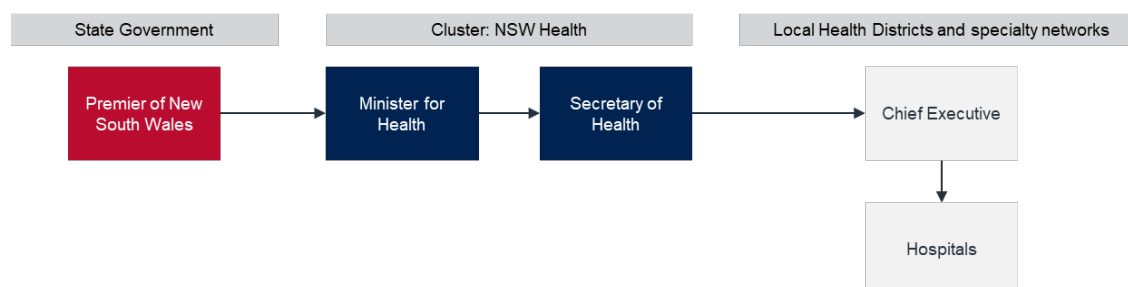
The NSW public health system is the largest public health system in Australia and falls under the supervision of the Ministry of Health. The Cluster is divided into 15 Local Health Districts (LHD) which comprise geographically contiguous groups of hospitals. There are 6 LHDs covering metropolitan regions and 9 covering rural and regional locations. Additionally, 3 Specialty Hospital Networks also operate alongside the LHDs. LHD's are headed by a Chief Executive who is supported by an Executive team.

LHDs are responsible for managing public hospitals and health services within their designated areas, ensuring the attainment and maintenance of adequate standards in patient care and services. While operating with a significant degree of autonomy, LHDs are ultimately accountable to the Minister for Health for their performance.

The NSW health system operates what it terms as a 'tight, loose, tight' model, whereby the Ministry enters into highly detailed service agreements with LHDs that cover service standards and funding allocation (tight), LHDs then have full autonomy over the use of this funding and their approach to meet service standards (loose), but face financial penalties and administrative sanctions from the Ministry should these standards not be met consistently (tight).

It should be noted that the Ministry rarely applies financial penalties and administrative sanctions on LHDs, even in circumstances where standards have not been met.

**Figure 10. Simplified delivery chain: NSW Health**



In 2015, the NSW government announced 12 Premier's Priorities for the period 2015-19, across a range of key areas important to citizens, including health. All priorities had specific and time-

bound targets for improved performance for citizen-focused outcomes. The Health System Priorities included:

- Improving service levels in hospitals, with 81 per cent of patients moved through emergency departments within four hours by 2019
- Tackling childhood obesity, by reducing overweight and obesity rates of children by five per cent by 2025

This case study will focus on NSW's experience of driving progress against the priority of improving service levels in hospitals across NSW.

To support the delivery of priorities, the Premier's Implementation Unit was set up in 2016. The Unit worked closely with frontline teams on data-led, evidence-based interventions to improve results. The Unit also routinely tracked and reported on progress against milestones and targets.

### **Target setting and measurement**

***The Ministry set an outcome focused, evidence-based hospital target, minimizing the need for consultation.***

'The four-hour rule', whereby a designated number of patients are seen within an Emergency Department within 4 hours is a well-established key performance indicator in the health system. Evidence shows that extended stays in hospitals, and in emergency departments in particular, lead to poor health outcomes for patients. Inefficiencies in Emergency Departments not only result in delays to the emergency treatment of patients, but also to the admission of more seriously ill patients who require specialized treatment in a non-emergency hospital setting. In 2015 the Ministry of Health set a target of moving 81% of patients through the emergency department within four hours by 2019.

Given the strong evidence base that underpinned the 'four-hour rule', consultation with LHDs was limited and was largely instituted top down. Chief Executives of LHDs were engaged but were provided with limited opportunity to meaningfully feed into the process of target selection.

***Capturing a holistic view of patient outcomes was critical.***

While there is evidence supporting the 'four-hour rule', the use of efficiency targets is a contested space amongst frontline clinical staff. Historically, the dominant culture amongst doctors, nurses and those on the frontline has been to 'treat the patient in front of you to the best of your ability', regardless of how long it takes. However, as the evidence builds for the causal link between efficiency and improved patient outcomes, this mindset is starting to shift.

Moreover, while the 'four-hour rule' target was the headline indicator for determining hospital performance, a number of other indicators were also tracked to ensure a holistic view on patient outcomes, as well as capturing potential unintended consequences that may result from single-mindedly pursuing the 'four-hour rule'. Some examples of these supporting indicators include rates of readmission and rates of mortality.



### ***Too many indicators can undermine focus.***

Targets have demonstrated effectiveness when used to cut through the malaise of complexity and to drive focus and urgency. A small set of targets to deliver hospital reform underpinned the Government's thinking behind the institution of the 'four-hour rule'.

However, the sheer number of KPIs that CEOs of LHDs were answerable to was overlooked. One CEO commented that at that time over 80 KPIs featured in his service level agreement with the Ministry, one of which was the 'four-hour rule'. Both administrators and frontline staff were continually questioning where to focus their efforts, and this would often change day-to-day. As a result, the 'four-hour rule' was often cast aside as part of the daily crisis management of hospital administration and clinical care.

### ***Tailored performance targets were set at a hospital level, which rolled up to the state target.***

Rather than adopting a 'one size fits all' approach to target setting, hospital level targets were set accounting for hospital size, scale of their operations and types of presentations (e.g., acute vs non-acute). Like hospitals were formed into four groups called 'Peer Groups', featuring:

- **Major Hospitals 1:** 35,000 patients and availability of one or more specialist services requiring specific infrastructure.
- **Major Hospitals 2:** 35,000 patients or less and no specialist services requiring specific infrastructure.
- **District Group 1:** 10,000 patients or less
- **District Group 2:** 4,000 patients or less

Targets were set for hospitals based on their Hospital Peer Group to account for their size and variation in the scale of operations. The State-wide target of moving 81% of patients through the emergency department within four hours was broken down as follows: 70% <4 hrs for Major Hospitals 1, 75% <4hrs for Major Hospitals 2, 85% <4hrs for District Group 1 and 95% <4 hrs for District Group 2. This approach allowed for contextualized targets, assigning higher targets to smaller hospitals who, due to their context, had greater efficiency. For larger hospitals, which had a more complex operating environment, lower targets were set.

Target setting was led by a central data and analytical function within the Ministry of Health to ensure a level of consistency, rigor and control over the process.

## **Performance Management Tools**

### ***Tailored support was provided to underperforming sites.***

The 'Whole of Hospital Team' was established in the Ministry of Health to provide targeted support to underperforming LHDs and hospitals. Featuring experts in a range of health disciplines, the team would embed in struggling hospitals, performing a diagnostic to determine the drivers of declining performance and then work intensively with the administration and clinicians to drive improvement.

The trigger for the Whole of Hospital Team's support was the Ministry's establishment of a 'Hospital Watchlist', whereby hospitals seeing a decline in performance for three consecutive months would be placed on a 'watchlist'. This involved the closer monitoring of these hospitals' performance and the involvement of the Whole of Hospital team in their improvement efforts.

***Making real-time data accessible to frontline clinicians played a key role in driving performance.***

The health system is steeped in data, but it is often not used effectively to inform day-to-day operations. Most hospitals in NSW have sophisticated data systems, with dedicated staff to service them. However, this function has been largely compartmentalized, both organizationally and physically, with the data team often situated in a separate wing of the hospital to clinicians. Moreover, real-time data was not made accessible to frontline clinicians and decision-makers limiting their ability drive improvements within their hospital.

As part of efforts to improve hospital efficiency, real-time data dashboards were designed and positioned in NSW's major emergency departments, enabling clinicians to make decisions in real-time that are informed by data. The dashboard was situated in a prominent location within the Emergency Department and featured critical information such as number of patients waiting, number of patients being treated, and real-time performance against the four hour rule.<sup>19</sup>

## **Incentives and consequences**

***The Ministry of Health's "tight-loose-tight" model has been largely ineffective at driving LHD accountability.***

As described above, service level agreements are in place between the Ministry of Health and each of NSW's 15 local health districts. These agreements outline a range of service standards that LHDs are expected to meet, and feature KPIs across clinical, financial, and administrative domains. In principle, LHDs who consistently fail to meet these standards are liable to financial penalties and the termination of key people in senior leadership roles (e.g., CEO). In practice however, this is rarely applied due to political factors and the fragile relationship that exists between the Ministry of Health and the frontline health system. As a result, while service level standards put LHDs on notice around their expected performance, there are minimal consequences in place should they fail to meet them.

***Data has been the Ministry's most effective mechanism in strengthening accountability.***

The availability of data at a granular level has enabled the Ministry to open dialogue with underperforming LHDs and hospitals about opportunities for support. There is a keen sense of awareness at all levels of the system that performance is being tracked, and that the Ministry will step-in to offer well-intentioned support when things are off-track. This creates a culture of high

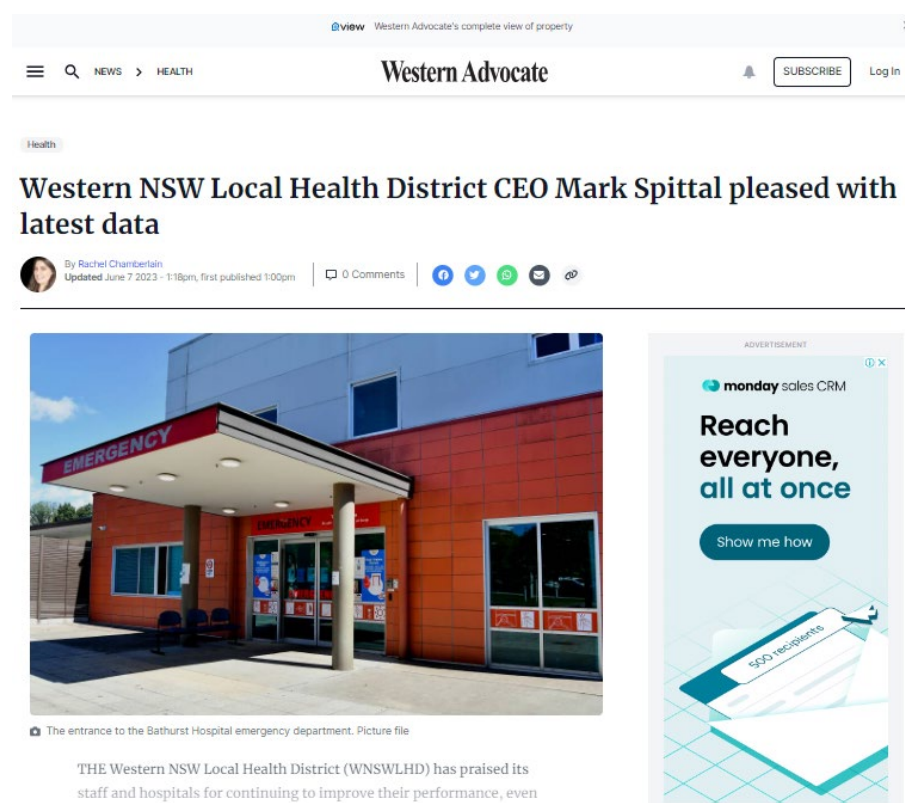
---

<sup>19</sup> An overview of NSW's Health EAV can be found at: <https://www.health.nsw.gov.au/wohp/Documents/EAV-user-guide.pdf>

performance without the overburdensome pressure that can often lead to perverse outcomes. This approach is facilitated by the timely access to granular level data.

Moreover, hospital performance against the 'four-hour rule' is released to the public every quarter by the Bureau of Health Information. This public release of data forges an informal contract between hospitals and the people they serve, whereby hospitals are held to account for quality and efficiency of service being offered. See Figure 12 below for an example of media coverage around the quarterly data release of hospital performance.

**Figure 12. Media coverage in one of the online newspapers in Australia**



***Like in Punjab, performance routines were also effective in strengthening accountability.***

The Premier of NSW, along with the Minister for Health, played key roles in holding the health system to account for achieving the '4 hour rule' target. Every month, a concise engaging 'note' was prepared by the Ministry of Health and the Department of Premier and Cabinet which outlined historical, current and projected monthly performance, as well as updates on the progress of key improvement strategies. These 'Monthly Notes' were keenly read by the Premier and Minister, with a hard copy often returned with comments and further questions jotted down in the margins.

Finally quarterly Stocktake meetings were held between the Premier, Minister for Health and Secretary of Health to conduct a deep-dive on hospital performance. Each of these meetings would end with a set of commitments made by the Secretary of Health to be achieved by the next

Stocktake. These meeting would open with a review of progress against each of these commitments.

These routines were highly effective at driving accountability within the NSW health system for three reasons:

1. There was a single person accountable for delivering on the reform, who led the quarterly meetings
2. The routines were data informed, they were not based on hearsay or anecdotes.
3. They were held frequently and consistently, they were rarely missed or rescheduled.

## Summary of key insights

### *What worked well*

- **Setting an outcome focused target:** The ‘four-hour rule’ target is underpinned by improving patient outcomes. This ensured that improvement strategies, financial investment and the prioritization of effort sought to achieve meaningful reform in the health system.
- **Establishing a hierarchy of targets:** Instituting hospital level targets accounting for their respective size and scale of operations ensured that the reform’s ambition was realistic at both a State-wide level, and on the frontline. Smaller hospitals which could achieve greater efficiency had higher targets than those larger hospitals that were operating in a more complex environment. Critically, these differentiated targets aggregated up to the State-wide target.
- **Tracking a diverse set of indicators:** While a headline target was set, supporting indicators were also tracked to provide a holistic view on patient outcomes and to also track any unintended consequences that may have resulted from achieving the ‘four hour rule’.
- **A focus on support:** LHDs and hospitals were provided with the support required to achieve their respective targets. With strong accountability came a commitment from the Ministry of Health to provide tailored assistance via the ‘Whole of Health Team’

### *What didn’t work well*

- **Public release of hospital performance data:** While the quarterly release of hospital performance against the ‘four-hour rule’ was effective at holding the health system to account, it also created immeasurable distractions. Following the data’s release, senior health leaders would inevitably be pulled in multiple directions such as managing media enquiries, answering to constituents, and carefully managing the Minister and his or her advisors. As a result, the focus on hospital performance was diminished over this period.
- **Too many KPIs in CEO service level agreements:** The sheer number of KPIs in CEO’s service level agreements made it increasingly difficult for LHDs and hospitals to prioritize

investment and improvement efforts. Often times, the health system was faced with multiple competing priorities without a clear sense of where the focus lay.



## 5 SOURCES

Barber, M., Rodriguez, N., Artis, E (2016). *Deliverology in Practice: How Education Leaders are Improving Student Outcomes*.

### Indonesia

*Achieving Quality of Educational Minimum Service Standards (MSS) through for Local Government*. <https://www.govtechedu.id/blog-product-updates/blog-post-rapor-pendidikan-daerah>

*Indeks Standar Pelayanan Minimal (SPM) Rapor Pendidikan Daerah*.  
<https://pusatinformasi.raporpendidikan.kemdikbud.go.id/hc/en-us/articles/22348260164889-Indeks-Standar-Pelayanan-Minimal-SPM-Rapor-Pendidikan-Daerah>

*Mengenal BOSP Kinerja Tahun 2023*. <https://ditsmp.kemdikbud.go.id/mengenal-bosp-kinerja-tahun-2023>

*Ministry of Education and Culture official webpage*  
<https://www.kemdikbud.go.id/main/blog/2023/08/lebih-dari-30-ribu-satuan-pendidikan-mendapat-dana-bosp-kinerja-kemajuan-terbaik-tahun-2023>

### NSW Education

*About Naplan Fact Sheet*. Retrieved from: [https://docs.acara.edu.au/resources/Fact\\_Sheet\\_-\\_About\\_NAPLAN.pdf](https://docs.acara.edu.au/resources/Fact_Sheet_-_About_NAPLAN.pdf)

*Learning from disruption: Why we should rethink the place of NAPLAN in our schools*, Monash University  
<https://lens.monash.edu/@education/2021/05/10/1383196/learning-from-disruption-why-we-should-rethink-the-place-of-naplan-in-our-schools>

*NAPLAN Review Final Report*. Retrieved from:  
[https://naplanreview.com.au/pdfs/2020\\_NAPLAN\\_review\\_final\\_report.pdf](https://naplanreview.com.au/pdfs/2020_NAPLAN_review_final_report.pdf)

*NSW Education official webpage*. <https://education.nsw.gov.au/>

### NSW Health

*Annual Report 2021-22 NSW Health (2022)*

*Local health districts and specialty networks*.  
<https://www.health.nsw.gov.au/lhd/Pages/default.aspx>

*NSW Health official webpage*. <https://www.health.nsw.gov.au>

*NSW Health Performance framework*.  
<https://www.health.nsw.gov.au/Performance/Pages/frameworks.aspx>

*NSW Service Agreement generic template 2023-24*. Retrieved from:  
<https://www.health.nsw.gov.au/Performance/Documents/service-agreement-generic.pdf>

## **Punjab**

Barber, Michael. *The good news from Pakistan* (2013)

*Transforming primary education in Punjab: the journey of Punjab's Education Reform Roadmap* (2018)