What is the relationship between politics, education reforms, and learning? Evidence from a new database and nine case studies

Unika Shrestha
Timothy P. Williams
Samer Al-Samarrai
Alies Van Geldermalsen
Attiya Zaidi

September 2019

Background Paper for World Development Report 2018

Abstract

Improving learning outcomes at scale is not just a matter of "scaling up." It also entails attending to the technical and political complexities that shape education reforms. In this study, we draw on country-level quantitative and qualitative data to study the systemic factors that contribute to improved learning or cause declines.

We approached this study through three activities. First, we established a database of education reforms and learning trends for 76 countries, enabling us to then examine the relationship between reform type and trends in learning outcomes. Second, using this database, we examined the relationship between education system characteristics, political and economic factors, and episodes of sustained improvements or regressions in learning. Third, we prepared case studies for nine countries from the database that helped show different themes concerning the politics of education quality reforms. We then synthesized our findings to identify trends in strategies for deploying information, working with coalitions, and creating opportunities for strategic change.

Findings from the database did not establish strong correlations with indicators of the economic and political conditions generally believed to have an effect on learning. Across most countries, the type of reforms introduced during periods of sustained learning improvements versus periods of decline did not differ systematically. Expansion in primary or secondary education access was not consistently correlated with trends in learning; however, lagged improvements in pre-primary enrollment was correlated with a greater probability of sustained learning improvements.

The findings also showed that changes in government spending on education was not strongly correlated with long-term learning trends. Moreover, neither economic growth during the episode nor lagged growth was correlated with learning trends. However, there was some suggestive relationship between accelerated growth and learning. In the case of political conditions, level of democracy failed to predict whether learning improved or declined. Episodes with the biggest improvements were more likely to begin in nondemocratic countries. However, there was some relationship between increased democratization (transition to a more democratic political regime) within countries and learning improvement.

Case study findings illustrated the complexities that shaped the ability of governments to deliver education reforms to improve quality. Cases of successful reforms and improved learning were characterized by governments whose decision-making and messaging were driven by information and learning metrics. These metrics used information to make the case for necessary reforms, provided incentives to improve local school systems and teacher quality, and established public accountability through a more informed citizenry. Accountability measures included international assessments, examinations, literacy rates, or school rankings. Coalitions and political incentives took different forms.

The case study findings also showed that the countries most effective in introducing and sustaining reforms considered the needs of various stakeholders at different levels of government and civil society. Those that failed to get the buy-in of a key group at the outset, e.g., teachers' unions, faced difficulties in implementing reforms, even if leaders were able to push through a policy reform. Effective communication strategy was also essential. The state needed to take control of the policy

reform's core message to prevent misinformation by competing interest groups. Effective reforms were focused and flexible. Policies were effective if they had a clear direction and could also be changed and even re-envisioned over time. Successful reforms were not necessarily contingent on charismatic leadership (though this could help). Sequencing popular reforms with those less likely to be supported helped to increase acceptance of less popular reforms. Reforms built on one another over a longer period of time, gradually adding greater levels of sophistication and nuance into the system in a way that slowly improved learning outcomes.

In conclusion, our study findings suggest that it is not any one component in education, economic system, or form of governance that is likely to improve learning. Rather, regardless of form, all technical inputs and political considerations must be coherent and aligned toward improved learning. Study findings and their implications must be read as exploratory. However, by offering some simple associations and hypothesizing the relationships of variables, the study offers a helpful perspective that can complement other recent efforts to understand the relationship between politics and quality reforms.

1. Introduction

In many developing countries, access to basic education has surged, but learning has not kept pace. Half of the 250 million children who cannot read, write, or do math go to school.[1] The lack of quality is concerning for many reasons. Education is considered a social good, a basic right, and a cornerstone for broader social and economic development.[2] Quality education has been shown to be a better indicator of economic growth than the number of years of school completed.[3] That so many children are in school but fail to acquire basic skills and competencies establishes a learning crisis, one with powerful socioeconomic and political dimensions and effects.

Ending the learning crisis will require effective use of the growing knowledge base on interventions that have improved learning. But improving learning outcomes at scale is not simply a matter of "scaling up." System-level technical and political factors can hamper scaling programs that may have worked at a smaller scale (for example, Kenya's experience of scaling up the use of contract teachers[4] and Cambodia expanding preschool coverage[5]). Thus, improving outcomes requires greater alignment of education stakeholders and learning goals as well as greater coherence between different elements of the education system. This will require governments to tackle the technical and political complexities of large-scale reform.

This study assesses the technical and political challenges that states face in their efforts to improve education quality. It investigates the system-level factors that contribute to improving learning and those that contribute to declines.

1.1. Research questions

In alignment with Theme 4 of the 2018 World Development Report (WDR), this study aims to answer the following questions:

- *Research question 1:* Are trends in learning outcomes correlated with the characteristics of education reform programs? This paper draws from a database of 76 countries to assess whether episodes of improved learning coincide with distinct types of education reforms compared to periods of declines in learning.

- *Research question 2*: Do broader economic and political factors differ systematically between episodes of improved learning and episodes of decline? It has been argued that education quality may be determined by factors such as economic growth,[6] education finance, institutional framework, and political conditions. This study explores whether the probability of experiencing episodes of improvement in learning differ by economic and political indicators before or during the learning episode.

[1] UNESCO (2014).
[2] Drèze and Sen (1995).
[3] Hanushek and Woessmann (2007).
[4] Bold et al. (2013).
[5] Bouguen et al. (2013).
[6] Bils and Klenow (2000).

- *Research question 3*: What are the technical and political challenges (and strengths) of education reforms, and what strategies have been used to overcome these challenges? Specifically, the analysis explores the role of information and knowledge, coalitions and incentives, and innovation in shaping education reforms.

1.2. Strategy and organization of the paper

We followed a three-step approach to answer our research questions. First, we identified episodes of improving or declining learning trends over five-plus years using panel data from cross-national standardized tests. Specifically, we used data for reading scores from the Program for International Student Assessment (PISA), the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), the Progress in International Reading Literacy Study (PIRLS), and the Latin American Laboratory for the Assessment of the Quality of Education (LLECE).

Second, we created a database of education reforms that were planned around the period of the identified learning episodes. Specifically, we documented types of reforms that were either planned or introduced during the period starting five years before the beginning of the learning episode until the last year of the episode.[7] In addition, we use data on economic growth, education spending, and political conditions corresponding to the learning episodes. This database helps answer our first two research questions.

Our third strategy included qualitative case studies on reform experiences for a diverse set of nine countries. The case studies also helped us dig into more nuanced political factors and reform processes that the variables in the database may not have captured. A primary criterion used for identifying a country for the case studies was the availability of learning outcomes and the ability to plausibly link these learning outcomes to reforms.

1.3. Summary of findings

Findings from the database do not establish strong and consistent correlations with indicators of economic and political conditions believed to affect education quality and learning outcomes. However, some patterns emerge:

- While there were many similarities in the education reforms associated with episodes of improving or declining learning in low- and middle-income countries, there were some notable differences. For example, curriculum, school-based management, learning assessments, and structural reforms were more common in episodes where learning had improved.

- There were also differences in the detail of broad reform efforts. While building institutional capacity was the most commonly introduced reform in low- and middle-income countries, activities to strengthen financial management systems were more common in episodes where learning improved.

[7] For example, for a learning episode from 2000 to 2007, we recorded programs that were planned or introduced between 1995 and 2007.

- Initial levels of gross domestic product (GDP) per capita and economic growth were not significantly different between episodes where learning had improved and episodes where it had declined.

- There was no strong evidence of a trade-off between improvements in learning and improvements in access.

- Spending per-student tended to rise more quickly in episodes where learning improved.

- Compared to episodes where learning outcomes deteriorated, democracy levels and the characteristics of a country's political settlement (an explicit or implicit agreement among powerful groups) did not seem to have a stronger association with episodes of improved learning.

- While there seems to be some relationship between the timing of reforms in the political cycle and improvements in learning, our information was insufficient to untangle whether successful reforms were introduced at the beginning or at the end of a cycle. However, government turnover was not more strongly associated with episodes of learning decline.

Findings from the case studies showed some of the complexities that shaped governments' abilities to deliver quality reforms.

- Elite buy-in was essential for a leader to introduce technical reforms. Successful reforms were not contingent on one charismatic leader (though this could help). Rather, quality reforms needed to be introduced by leaders who were informed by well-qualified experts.

- Governments committed to improving learning outcomes tended to have information drive their decision-making. Information and metrics helped build demand for reforms that could track learning, improve the curriculum, assess performance, and so on.

- Information could also drive communication strategies, be used to make the argument for reforms, and gain the support of key actors, including the public. An effective communication strategy could also counter the spread of misinformation by groups who oppose the reform.

- Effective coalitions enlisted the buy-in of stakeholders that would be vital for introducing and implementing the reform. Teachers' unions were often a powerful opposition group. Taking a confrontational stance to fight unions ran the risk of alienating the group of teachers who would implement the reform. In the country cases we looked at, teachers' unions often succeeded in obstructing efforts to introduce quality reforms.

- The most effective reforms were focused but flexible. Policies were effective if they had a clear objective, but they could also be changed or re-envisioned as external conditions

changed. Over a longer period, reforms could build on one another, gradually offering greater levels of sophistication to the system in a way that could incrementally improve learning.

- "Critical moments" could give rise to quality-focused reforms. Leaders could feel compelled to accept reports if there is, for example, a change in leadership, a major political shift in the country (e.g., end of apartheid), or a report highlighting low learning outcomes.

- Sequencing popular reforms with those less likely to be supported helped increase leaders accept less popular reforms. In the case of teacher accountability, starting with collective incentives could have helped build an evaluation culture that could have made it easier to implement individual incentives later on. Such an approach was more likely to guarantee leaders implement meaningful reforms rather than launch a drastic policy change that may be rejected from the outset.

The next section reviews literature on some of the technical and political challenges that governments face when introducing education quality reforms and improving learning outcomes.

2. Literature review: The technical and political challenges of improving learning outcomes

Education systems and learning outcomes are shaped by a range of technical and political complexities.[8] Achieving systemic change requires aligning these various components toward student learning. But even when evidence points to clear interventions that will improve learning outcomes, it is not always acted upon.

Education systems often are constrained by several technical challenges that limit their effectiveness. First, financing for education systems is often insufficient. Education systems and policies may be well designed, but without adequate levels of funding, even the best designed interventions will not work. Yet most education systems do not yet spend the recommended 6% of GDP on education.[9]

Second, available resources in education systems are often not well used. Teachers may not be prepared to be effective in their work. Many lack enough training and pedagogical skills to deliver the curriculum. Salaries or other benefits may be low, leaving teachers with little incentive to show up for their own classes. They may take on second jobs outside of teaching, even moonlighting as private tutors, sometimes for their own students.[10]

Third, linking inputs to learning outcomes is not straightforward. Many states have proven adept at building classrooms, increasing access, offering trainings, or introducing new materials to improve pedagogy. However, the mechanism that links these important investments with students' acquisition of basic skills and proficiencies is poorly understood.[11] Even conventional technical inputs such as textbooks may not bring about the expected improvements to children's learning.[12]

The challenges of improving quality also face political complexities.[13] Scholars have pointed out that it is reasonably easy for leaders in democratic states to make a case for improving access because these activities are straightforward, visible, and popular.[14] But improving education quality is more challenging than building classrooms, as it is often less visible and can be more complex. The evidence base for how to improve learning outcomes is underdeveloped, particularly in resource-limited settings. Even when promising interventions are identified, results may not come quickly enough for leaders, who often need to determine the political implications of investing in quality education alongside their voters' other interests.[15]

---

[8] Many of the arguments in this section were informed by the work of Hossain et al. (2017), Hickey and Hossain (2019), and Bruns, Macdonald, and Schneider (2019).
[9] UNESCO (2014).
[10] Bray and Lykins (2012); CAMPE (2015).
[11] Pritchett (2013).
[12] Ulrich, Kremer, and Moulin (2009).
[13] Grindle (2004); Hanushek and Woessmann (2007); Kingdon et al. (2014); Kosack (2012); Pritchett (2013).
[14] Hossain and Moore (2002); Mani and Mukand (2007); Grindle (2004).
[15] Kosack (2012).

In short, the incentive for political leaders to introduce reforms to improve education quality may not be as politically advantageous as we might assume. Quality-focused reforms can entail fighting powerful teachers' unions, losing jobs, or shifting budget priorities away from higher education.[16] Each of these factors can mean losing elections.

The elite have formal and informal coalitions, whose priorities and incentives may or may not align with a commitment to equity and quality. They have other interests, such as maintaining power and allegiance to the "vital constituencies"[17] that are central to keeping them in power or championing the needs of special interest groups. Thus, education policies may *or may not* be introduced with the expectation that learning outcomes will improve.[18]

Particularly in weaker states, there may be less pressure exerted on governments to improve quality.[19] Public forms of accountability cannot be assumed, and demand for education workers may change. In less developed settings, businesses and industries may have little need for an educated workforce, placing the burden for improving learning on the state and nongovernmental organizations (NGOs).

Demand may not be strong from households, either. Parents themselves may lack a formal education, so they may not know what to demand or expect from their government when it comes to quality.[20] Without demand from businesses or households, the most powerful source of demand to improve quality may have to come from the state itself. This demand must align with global commitments concerning sustainable and equitable development for all and with the global movement to promote mass education.[21]

The politics of education policies and learning outcomes can be studied at different levels. At the national level, there is a need to understand the relationship between learning outcomes and the national political settlement.[22] The priorities and incentives for introducing and implementing pro-poor policies can be studied by looking at the individuals, incentives, and institutions that would lead the national government to prioritize quality and equitable education.[23]

At the education sector level, the study of politics can focus on the incentives and forms of accountability of local government officials, education officers, and teachers. For example, it can focus on the effectiveness of decentralization or on the impact of teachers' unions with respect to learning outcomes, examination performance, or completion rates.[24]

---

[16] Bruns, Macdonald, and Schneider (2019).
[17] Kosack (2012).
[18] Kingdon et al. (2014); Kosack (2012); Grindle (2004).
[19] Harding and Stasavage (2013).
[20] Kosack (2012).
[21] Boli, Ramirez, and Meyer (1985).
[22] Political settlement refers to "the balance or distribution of power between contending social groups and social classes, on which any state is based," according to Di John and Putzel (2009, p. 4).
[23] Khan (2010); Kingdon et al. (2014); Williams (2017).
[24] McLoughlin (2011); Williams (2019).

An investigation into the politics of education quality must also examine the role of informal politics.[25] Who are the actors and interest groups propelling the trajectory of the sector, and why might they support quality and equitable education policies? Approaching a study of education politics using this lens can help us understand the distribution of power between different social and political groups that shape the implementation of education reforms. In short, the study of politics and governance in education can help reveal patterns associated with accountability, technocratic effectiveness, forms of service delivery, and the alignment between national-level goals and incentives to improve learning.

We now turn to examine findings from the cross-country database to examine episodes and trends of education reform and learning improvement.

[25] Grindle (2004); Kosack (2012).

3. Episodes of learning improvement using cross-country data

In this section we aim to explore the education, economic, and political contexts associated with successful episodes of learning improvement by drawing together the available cross-country data. We first use international and regional assessments to identify country episodes where learning outcomes changed. Using these "learning episodes" as our main unit of observation, we then build a database of relevant education, economic, and political variables associated with these episodes. We use this database to compare and contrast the characteristics of country episodes of improving or deteriorating learning outcomes. The analysis in this section does not identify causal determinants of changes in learning but instead explores factors correlated with periods of sustained learning improvements versus periods of decline.

3.1. Approach

Episodes of learning improvements (or declines) were identified using data on reading assessments from the PISA, LLECE, SACMEQ, and PIRLS assessments. We selected reading because the PISA assessment covers a large number of countries and reading scores are reported on a common scale over a long period of time.[26] Moreover, more data are available for reading in developing countries than for math and science. For example, in the first round of the SACMEQ tests conducted between 1995 to 1999, students were only tested in reading.[27]

To identify periods of changes in learning, we adopted a standard set of rules across all countries and assessments (see Box 1 and Annex A). In total, 105 episodes across 76 countries were identified.

Box 1. Identifying episodes of changes in learning

Using information on learning outcomes, we established a series of learning episodes based on intervals of a minimum of five years, for each country. These episodes identify the main direction of change in learning outcomes and the period that these changes occurred in. For example, using the Progress in International Reading Literacy Study (PIRLs) data, we identified learning changes between 2001 and 2006 and between 2006 and 2011. In the case of the PISA data, we checked the direction of change in learning for six-year periods between 2000 and 2006, between 2003 and 2009, and for all other periods between 2000 and 2015.

Once these periods were identified, we then sought to establish the direction of change in learning and whether these changes were statistically significant. The periods were classified as improving/worsening if the change in learning was statistically significant. Periods that had statistically insignificant changes in learning were coded as "no change" episodes and were excluded from our database.

Since each interval used in the PISA assessments included data from three test years, it is possible that learning outcomes did not change monotonically. We only included periods where the change between the first and final year was statistically significant.

Once each period was categorized as an increasing or declining period, we extended the overall period they covered to identify learning episodes based on the following rules.

[26] Results of reading assessments from all six rounds of the PISA tests have been reported on the same scale. That is not the case for results of math or science assessments; see OECD (2016).
[27] See SACMEQ (2018).

Let *k* denote the last year of the interval and *j* denote the first year of the interval. For an increasing interval,

- if the score in the assessment year immediately after the end of the interval (year *k+1*) was greater than the score at the end of the original period, then the period was extended to include year *k+1*; and
- if scores remained unchanged or decreased between year *k* and year *k+1*, but the change was not statistically significant, then the episode was expanded to year *k+1* as long as the score in year *k+1* was significantly large than the score at the beginning of the episode (year *j*).

We applied these steps to the second and third test years following the end of the original interval.

- If the score in the assessment year preceding the beginning of the interval (year *j–1*) was lower than the score in year *j*, then the interval was extended to include year *j–1*.
- If scores remained unchanged or decreased between year *j–1* and year *j*, but the change was not significant, then the interval was expanded to year *j–1* if the score was significantly lower than the score in the end of the new interval. These steps were applied to the second and third years preceding the beginning of the original interval.

We followed analogous rules for extension in the case of intervals coded as declines (further details are available from the authors).

In determining learning episodes, the PISA tests were given precedence. For countries that participated in the PISA, as well as in one of the other three assessments (e.g., Bulgaria, Chile), we identified learning episodes based on PISA scores. If a country participated in the PISA tests but did not change scores significantly over the years, we still included them if scores from another assessment showed significant trends. In the case of all assessments, we excluded from the database countries that participated only once. In addition, we also excluded countries in the PISA tests that did not have at least two data points that were six years apart.

3.2. Reforms database construction and description

We collected information on the education reforms that were taking place at the time of each learning episode from two main sources:

- United Nations Education, Scientific, and Cultural Organization's (UNESCO) Planiopolis database provided information on national policy documents for many countries. These documents included National Education Plans, National Education Legislations, and Education for All reports.[28]
- Eurydice, an online database of legislation filled gaps in documentation for European countries.

Reforms included programs or policies planned in the period that began five years before the start of a learning episode and throughout the episode itself (hereafter referred to as "the reform period").

Reforms were coded using two main categories. First, we categorized each program or policy mentioned in the sources into one of 15 broad **topics**. These topics ranged from policies related to improving teacher quality to structural reforms of the education system. Reforms were further coded into **activities** within each broad topic. For example, a program related to developing the information system would be classified under the topic of "building institutional capacity" and the

[28] For some countries, documents were unavailable in English and were not included in our search.

activity of "development of management information systems."[29] In addition, the education subsector covered by the reforms is also documented when this information was available.

The sources we used to capture the reforms occurring during our learning episodes only include information on planned policies and interventions. We have no consistent information on whether planned policies were implemented or whether implementation was consistent with the original plans.

In addition to variables indicating the period and trends in learning and the type of reforms undertaken over the period, the database also includes information on the concurrent economic and political context. Table 1 indicates the additional variables and sources included in the database.

Table 1: List of indicators and original data source

Indicators	Original source	Number of observations
Growth in enrollment rate during episode, primary	UNESCO Institute of Statistics (2017)	59
Growth in enrollment rate during episode, secondary	UNESCO Institute of Statistics (2017)	51
Growth in pre-primary enrollment (lagged)	UNESCO Institute of Statistics (2017)	80
Spending per student	UNESCO Institute of Statistics (2017)	
Timing of executive elections	IADB-Database of Political Institutions (2015)	64
Tenure of administration	IADB-Database of Political Institutions (2015)	97
Number of executives	IADB-Database of Political Institutions (2015)	76
Degree of fractionalization	IADB-Database of Political Institutions (2015)	96
GDP per capita (in constant PPP dollars)	World Development Indicators	103
Government effectiveness	World Governance Indicators	101
Polity score, 2013	Polity IV Project	96
Type of political settlement in education, 2000	Levy (2014)	56

Notes: IADB stands for Inter-American Development Bank, UNESCO for United Nations Education, Scientific, and Cultural Organization, and PPP for purchasing power parity.

The database includes information on learning outcomes, education reforms, and economic and political conditions for a total of 105 episodes across 76 countries. Since the PISA test has the greatest number of participating countries, 73% of the learning episodes in the database are based on PISA reading scores, while the share of episodes based on SACMEQ reading scores is 14%. Episodes based on the LLECE and PIRLS reading tests comprise 6% each of all the episodes. Since the PISA only test targets secondary students, most (73%) episodes are based on secondary level learning outcomes. In terms of learning trends, 61% of episodes represent improvements, while 39% represent declines in learning.

A shortcoming of our database is that low- and lower-middle-income countries are underrepresented because they are less likely to participate in cross-national assessments.[30] As Table 2 shows, high-income countries make up 36% of all countries but account for over 50% of

[29] A list of the topic and activity categories are available on request.
[30] We use the World Bank's income classification as of 2017.

countries included in our database. European and Central Asian countries are overrepresented in our database, while sub-Saharan African countries are underrepresented, and there are no South Asian countries.

Comparing economic and political characteristics shows that countries in the database are like the universe of countries within their income group except in the case of polity scores. Average polity scores suggest that countries in the database are more democratic (i.e., they have higher scores, on average).

Table 2: Overall and database distribution of countries by various characteristics

Panel A	**Overall (%)**	**Database (%)**
Region		
East Asia & Pacific	16.9	7.9
Europe & Central Asia	25.6	47.4
Latin America & Caribbean	19.8	19.7
Middle East & North Africa	9.7	7.9
North America	1.5	2.6
South Asia	3.9	0
Sub-Saharan Africa	22.7	14.5
Number of countries/episodes	207	76
Income classification		
Low income	14.5	4
Lower-middle income	23.7	11.8
Upper-middle income	26.1	30.3
High income	35.8	54
Number of countries/episodes	207	76
Panel B		
Per-capita GDP (constant, PPP) by income group		
Low & lower-middle income	4,009	5,386
Upper-middle income	13,862	14,872
High income	41,903	39,328
Number of countries/episodes	177	75
Polity score, 2013 by income group		
Low & lower-middle income	2.5	4.1
Upper-middle income	**3.7**	**6.8**
High income	**6.9**	**9.1**
Number of countries/episodes	159	72
Per-student spending: Primary, 2012–14		
Low & lower-middle income	560	962
Upper middle income	2,084	2,272
High income	7,552	7,723
Number of countries/episodes	91	49
Per-student spending: Secondary, 2012–14		
Low & lower-middle income	816	1,143
Upper-middle income	2,481	2,601
High income	8,889	8,503
Number of countries/episodes	89	47
Gross enrollment: Primary, 2013–14		
Low & lower-middle income	105	109
Upper-middle income	106	107
High income	104	103
Number of countries/episodes	157	69
Gross enrollment: Secondary, 2013–14		
Low & lower-middle income	60	59

Upper-middle income	93	96
High income	107	111
Number of countries/episodes	140	65

Notes: Numbers in bold denote statistically significantly different averages. PPP stands for purchasing power parity.

3.3. Are improvements in learning associated with specific types of reforms?

There are many similarities in the broad types of education reforms (associated with improvements or declines in learning) pursued in country episodes (Table 3). In low- and middle-income country episodes, the most common reforms planned were those related to system strengthening, efforts to improve teacher quality, and interventions aimed at addressing education inequalities. A broadly similar share of episodes of improving and declining learning outcomes included these types of reforms. However, compared to episodes where learning outcomes declined, there are some striking differences in some of the reforms associated with episodes of improved learning and decline.

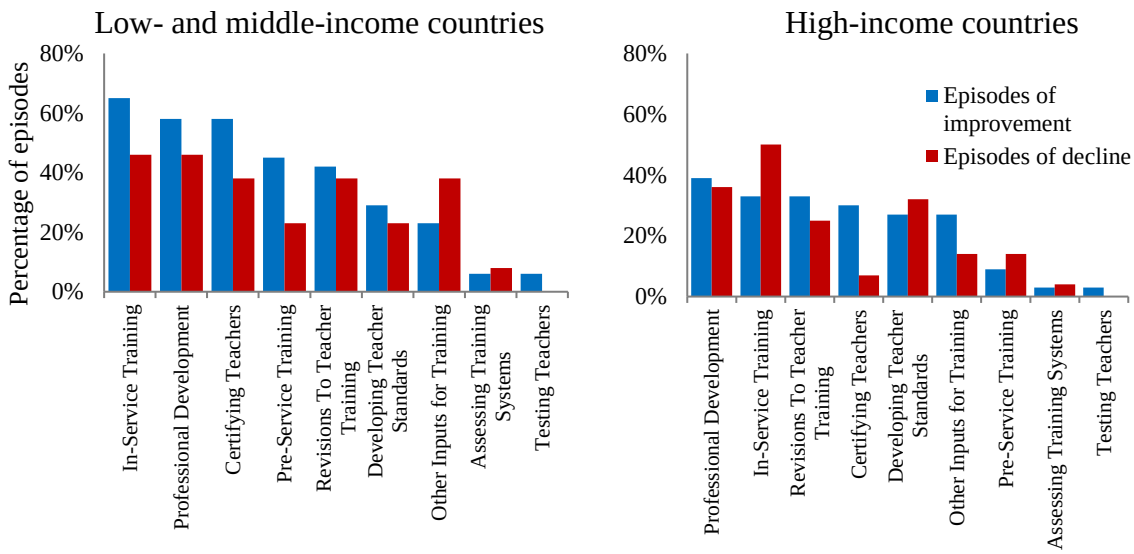
Table 3: Share of episodes that include specific reforms by direction of change in learning

Broad reform area	**Low- and middle-income countries**		**High-income countries**	
	Share of episodes where learning has		**Share of episodes where learning has**	
	improved	**declined**	**improved**	**declined**
Improving teacher quality	97	85	82	82
Ensuring equity in access and retention	94	77	94	82
Building institutional capacity	94	85	64	57
Curriculum reforms	87	69	91	86
School-based management	87	38	70	79
Learning assessment systems	84	62	70	79
Structural reforms	84	62	73	71
Provision of textbooks, learning materials,	84	69	42	39
skills development	74	54	67	68
ICT use in education	71	46	73	36
School/classroom construction for expanding access	71	62	24	25
Teacher pay-and-performance evaluations	39	15	21	29
Hiring more teachers	23	0	21	18
Other approaches to learning (besides ICT use)	13	0	15	14

Notes: The most common types of reforms were similar between episodes of learning improvement and periods of decline. ICT stands for information and communications technology.

However, looking in more detail at the types of activities under each broad reform type reveals differences between episodes of improvement and decline. For example, Figure 1 shows that within the reform category of improving teacher quality, reforms to certify teachers were more common during episodes of learning improvement. In low- and middle-income countries, 58% of improvement episodes were associated with teacher certification, compared with only 38% of declining episodes.

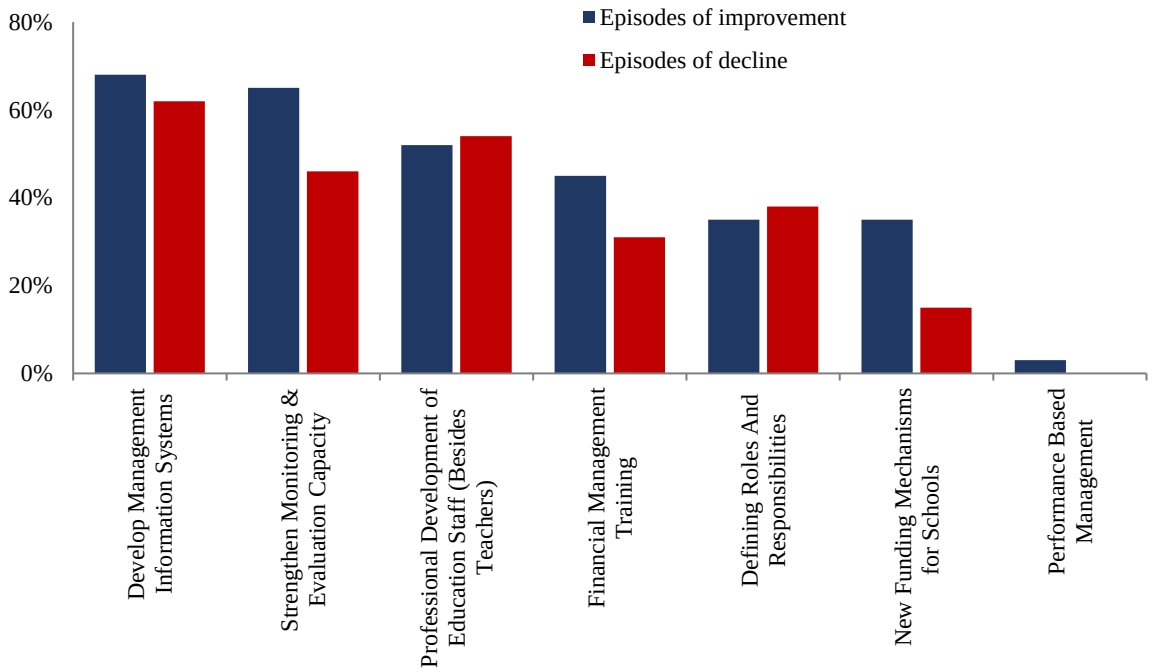
Figure 1: Share of episodes that included one or more reform activity



Notes: The figure shows that teacher certification reforms were more likely to occur during episodes of improvement.

The share of improvement episodes during which one or more reforms related to strengthening financial management capacity was 45%, compared to only 31% in episodes where learning declined (Figure 2). Episodes of improvement were also more likely to include the introduction of new funding mechanisms for schools. Measures to build monitoring and evaluation capacity were also more likely during episodes of improvement. The share of improvement episodes that coincided with reforms to strengthen monitoring and evaluation was 65%, whereas such reforms were introduced during 46% of episodes of decline.

Figure 2: Share of episodes that included one or more reform activity, LICs, and MICs



Notes: The figure shows that reforms to strengthen financial management and accountability were more likely during episodes of improved learning. LIC stands for low-income country and MIC for middle-income country.

3.4. Are improvements in learning associated with other education system conditions?

Changes in other elements of the education system may also affect learning outcomes. For example, rapidly increasing enrollment without a commensurate adjustment in inputs or in practices to maintain good-quality education delivery could compromise learning. Conversely, a drop in enrollment may reflect unfavorable conditions in the broader education system that may also be correlated with learning. Increased focus on access could also prolong low attention to learning since expanding access is often a more politically favored cause than improving learning. Investment in early childhood education, changes in public expenditure on education, and factors like teacher status could also affect learning outcomes.

In the database, the association between changes in enrollment and changes in learning differ by level of education, but they are generally weak and not statistically significant. Our learning episodes drawn from the PISA measure learning at the secondary school level. For episodes of learning improvement, secondary enrollment ratios increased by about 5 percentage points (pp) compared to only 3 percentage points in periods where learning declined (Table 4).[31] This does not support the notion that there is a quantity-quality trade-off in education where increases in enrollment necessarily lead to reductions in learning.

[31] Australia and Sweden are included as declining learning episodes. They experienced unusually large drops in secondary enrollment, over 25 percentage points.

Our other learning episodes mostly relate to changes in learning outcomes at the primary level. In this case, episodes where learning outcomes improved were associated with smaller increases in enrollment ratios. This relationship is strongest among low- and middle-income countries (see Table 4) and suggests there may be a trade-off between increasing access and improving learning at the primary level. However, these differences are not statistically significant.

Table 4: Change in gross enrollment rates by episode type

	Indicator for episodes where learning has	
	improved	**declined**
*Average percentage point change in GER***:**		
Secondary (PISA) level-based learning episodes (all countries)	5.4	2.9
High-income countries	3.6	2.5
Low- and middle-income countries[32]	9.8	5.4
Primary level based learning episodes (all countries)	2.5	4.3
High-income countries	–0.84	–0.88
Low- and middle-income countries	2.99	6.83

Notes: The table shows there is no large or statistically significant trade-off between expanding access and improving learning outcomes. PISA stands for Program for International Student Assessment and GER for gross enrollment ratio.

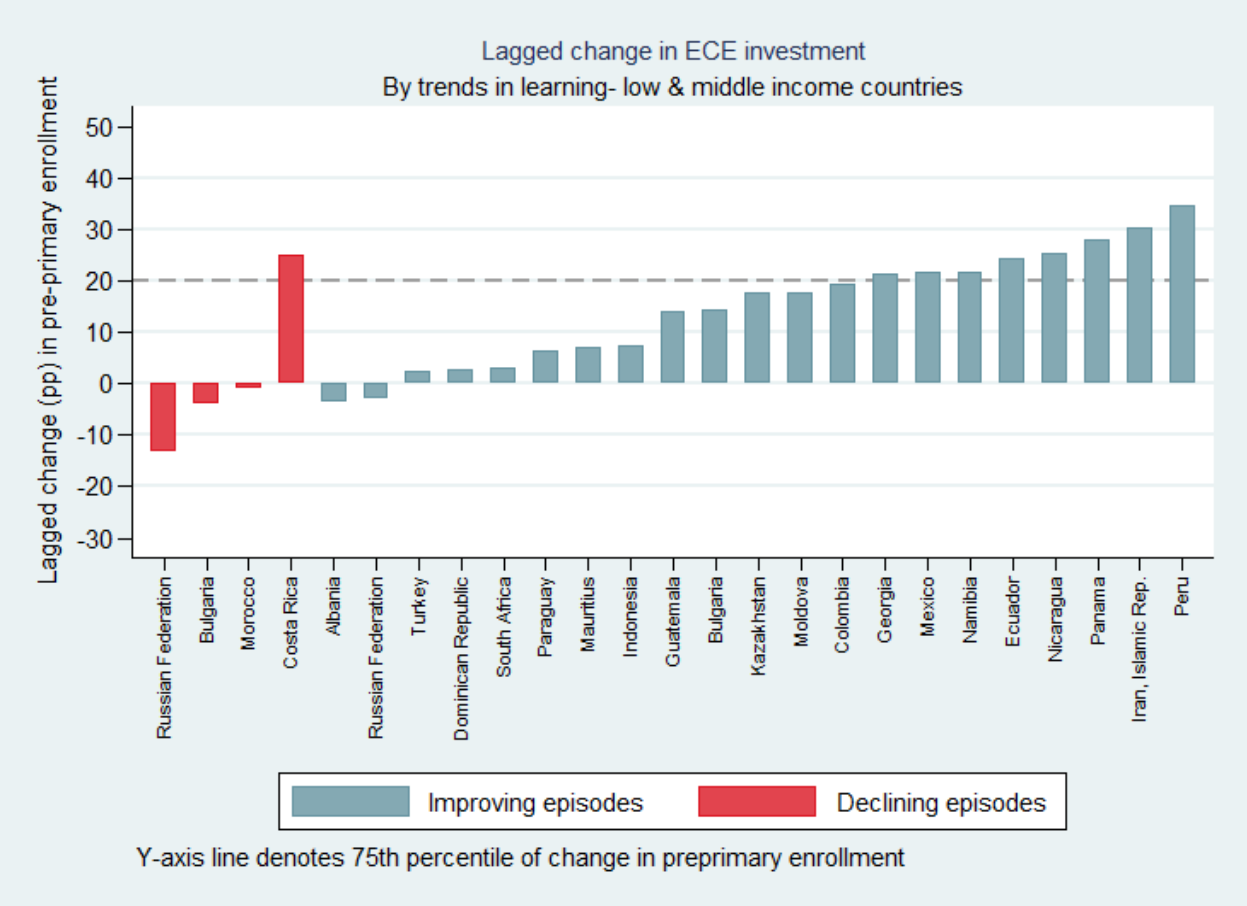
Based on lagged changes in pre-primary enrollment, the data suggest an association between investments in early childhood education and improved learning outcomes. Foundational skills acquired in the early years can be central to improving future cognitive skills. To account for the lag in the impact of investments in early childhood education on primary and secondary learning outcomes, we looked at changes in pre-primary school access that occurred at least 7–10 years before each learning episode in our database.[33] Countries with episodes of learning improvement experienced a bigger increase in lagged pre-primary enrollment compared to countries with learning declines (Figure 3). In low- and middle-income countries, the average change in lagged pre-primary enrollment was 15 percentage points during episodes of improvement, while the change was 2 percentage points during periods of decline. This difference is statistically significant at the 10% level and does not seem to be driven by a particular income group.[34]

[32] This is based on only three episodes.

[33] For example, we compare the change in pre-primary enrolment over a period of the same length as the learning episode but ten years before the episode's first year for the PISA-based episodes and seven years before the episode's first year for primary level assessments.

[34] We note, though, that this analysis doesn't establish causality.

Figure 3: Lagged change in ECE enrollments by trends in learning: Low- and middle-income countries



Notes: The figure shows that larger increases in early childhood education enrollments are associated with episodes where learning improved. Countries are ranked along the horizontal axis in order of the change in pre-primary enrollment. The y-axis line denotes 75th percentile of change in pre-primary enrollment. PP stands for percentage points and ECE for early childhood education.

Government spending on education, if invested in the right areas, has the potential to improve learning. Taking all countries together, episodes of learning improvement tended to be more common in countries with lower initial levels of spending per student (Table 5). And spending per student rose more rapidly during periods of learning improvement compared with episodes of decline. But looking at different income groups yields a highly varied picture (recognizing that sample sizes are quite limited). At the primary level, there is positive association between improvements in learning and improvements in spending per student in low- and middle-income countries. At the secondary level, spending increases are similar across episodes where learning increased or decreased.

Table 5: Average annual change in spending per student

	Low- and middle-income countries		High-income countries		All countries	
	Indicator for episodes where learning has		Indicator for episodes where learning has		Indicator for episodes where learning has	
Spending per student	**improved**	**declined**	**improved**	**declined**	**improved**	**declined**
Government spending per primary school student:						
At the start of the learning episode	875	487	10,265	6,763	2,049	2,579
Annual % change during learning episode	5.2	–1	–1	2	4.4	–0.03
Number of episodes	14	2	2	1	16	3
Government spending per secondary school student:						
At the start of the learning episode	1,758	2,246	7,469	8,415	6,041	7,879
Annual % change during learning episode	6	7.4	2.5	2.4	3.4	2.8
Number of episodes	8	2	24	21	32	23

Notes: Spending averages are provided for primary and secondary education levels depending on the level that the learning episodes relate to. For example, averages for the PISA (Program for International Student Assessment) learning episodes are for spending per secondary school student.

While the scant coverage of data on spending by education level makes it difficult to draw firm conclusions on the association between changes in learning, we can compare salary spending as a share of total public education spending. Based on available data for 57 episodes, we find that staff compensation as a share of total education spending decreased more during episodes of improvement than during episodes of decline.[35] Put another way, the share of nonsalary spending increased more during episodes of improvement than during episodes of decline. However, this difference is not statistically significant and seems largely driven by three countries—Peru, Chile, and the Dominican Republic—where the share of staff pay decreased by over 20 percentage points during their respective learning episodes.

3.5. Are increases in learning associated with economic conditions?

Favorable economic conditions may contribute to improvements in learning outcomes. For example, countries with higher levels of national income and government revenues may have more resources to support quality-enhancing education reforms. They can also have higher institutional capabilities to plan, design, and implement reforms. Higher levels of per-capita income can also lead to better learning outcomes through improving health and family investments in children and education.[36]

[35] Episodes of improvement coincided with an average decrease in the share of staff pay of 3.4 pp, while the average decrease among episodes of learning decline was –0.6 pp.
[36] Bruns and Luque (2014); Filmer and Pritchett (1999).

Comparing levels of GDP per capita between episodes of learning improvement and decline shows no statistically significant difference.[37] However, if the learning episodes in low- and middle-income countries are ranked according to the size of the change in learning, the top 20% of country episodes where learning improved had higher levels of GDP per capita than the bottom 20% of country episodes where learning declined.[38]

Faster economic growth also does not seem to be associated with episodes of improved learning. Across income groups, the average annual growth rate during the first five years of the episode is not significantly correlated with whether learning improved or declined during that episode. Even allowing for a lag in the effect of growth on learning trends does not show a statistically significant association.

Table 6: GDP and GDP per-capita growth by type of learning episode (constant PPP dollars)

	Low- and middle-income countries		High-income countries		All countries	
	Indicator for episodes where learning has		Indicator for episodes where learning has		Indicator for episodes where learning has	
	improved	declined	improved	declined	improved	declined
Average GDP per capita at start of episode	9,078	7,031	33,149	33,695	21,890	26,174
Average growth in GDP per capita in first 5 years of episode	3.7	3.3	1.3	1.6	2.4	2.1
Average growth in GDP per capita in 5 years before episode	2.8	2.4	3.0	3.0	2.9	2.8

Notes: The table shows that the level and growth of GDP per capita is not associated with changes in learning. GDP stands for gross national product and PPP for purchasing power parity.

3.6. Are increases in learning associated with political conditions?

While studies have associated democratization with an expansion in education provision,[39] our study did not find a strong link between democratization and learning outcomes. We used polity scores at the beginning of each learning episode to measure the level of democracy in the database. Our analysis shows that the level of democracy is not strongly associated with the direction of change for the learning episodes (Table 7).

Comparing changes in the democracy measure within countries over the course of the episode by learning trends may be a more useful exercise. Although few episodes coincided with a change in

[37] GDP per capita is measured at the start of the learning episodes.
[38] Ranking country episodes in this way for high-income countries shows no statistically significant difference.
[39] Harding and Stasavage (2013); Acemoglu et al. (2014); Ansell (2010); Stasavage (2005).

polity scores, periods of improved learning tended to be associated with smaller improvements in polity scores than episodes of decline (Table 7).[40]

Table 7: Polity scores by type of learning episode

	Low- and middle-income countries		High-income countries		All countries	
	Indicator for episodes where learning has		Indicator for episodes where learning has		Indicator for episodes where learning has	
	improved	declined	improved	declined	improved	declined
Average polity score at start of learning episode	5.6	4.8	8.9	9.7	7.3	8.2
Average change in polity score during learning episode	0.1	0.5	0.03	–0.08	0.07	0.09

Notes: The table shows that levels of democracy are not associated with changes in learning. Polity scores range from –10 to +10, with –10 corresponding to complete autocracy and 10 to complete democracy.

The political settlement under which a country operates could affect whether education reform programs improve learning outcomes. For example, settlements with high levels of clientelism[41] may not be particularly conducive to improvements in learning since stakeholders in the education system may be more motivated by other objectives in that system. Levy[42] classifies countries into three different categories based on the concentration of authority: dominant settlements have concentrated power and high levels of elite cohesion and clientelism, while competitive settlements have lower levels of elite cohesion but less clientelism.[43]

The type of political settlement was not disproportionately associated with either episodes of learning improvement or decline (Table 8). However, the average decrease in reading scores among countries that had highly concentrated power in 2000 was much larger than for countries with a more competitive settlement. However, this is based on merely two countries.[44] Moreover, these countries also experienced larger increases in scores during episodes of improvement. Therefore, it is difficult to draw definitive conclusions. The fact that indicators of the type of political settlement are only available for 2000 presents further difficulties in interpreting the results.

[40] A fifth of the episodes with available data on polity scores experienced a change in polity scores between the beginning and end of the episode.
[41] Clientelism is when political parties reward supporters with benefits that have been funded by taxpayers.
[42] Levy (2014).
[43] These indicators are only available for 2000 and for only 56 learning episodes.
[44] Morocco and Tunisia.

Table 8: Type of political settlement and its effect on learning trends

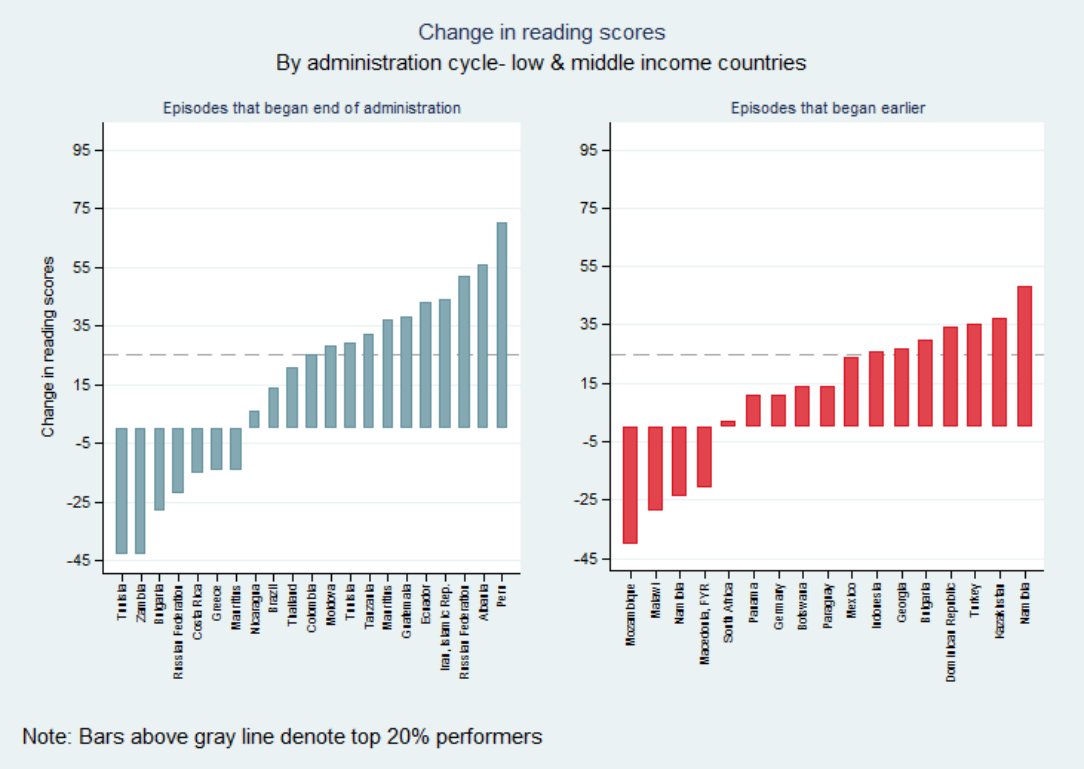
	Low- and middle-income countries			All countries		
	Declining episodes	**Improving episodes**	**Overall**	**Declining episodes**	**Improving episodes**	**Overall**
Percentage of episodes from countries by political settlement type, 2000						
Competitive	18.2	28.6	25.6	43.8	45.0	44.6
Intermediate	63.6	53.6	56.4	43.8	42.5	42.9
Dominant	18.2	17.9	18.0	12.5	12.5	12.5
Average change in reading scores (points)						
Competitive	–15	23	15	–21	21	9
Intermediate	–30	33	13	–30	30	11
Dominant	–41	32	11	–41	32	13
Number of episodes	11	28	39	16	40	56

Notes: High-income countries are not reported separately because there are too few cases.

Even though democracy levels do not seem to be strongly correlated with learning trends, we looked at when learning improvements take off. Data on the timing of executive elections suggest that episodes of improvement were more likely to begin close to an election year. For 75% of improving episodes, an executive election occurred within the first three years of the episode, compared with 58% among declining episodes. This difference, however, is not statistically significant.

Episodes with the largest improvements in learning tended to begin at the end of an administrative cycle. Figure 4 shows the magnitude of proportional changes in reading scores by the time they started in an administrative cycle. The blue bars show the size of changes in learning for episodes that began when the chief executive had one year or less remaining in his or her term. The red bars show episodes that began earlier in the administration. Sixty-two percent of the largest improving episodes (denoted by bars above the gray line in Figure 4) began at the end of the administration, compared to 41% of other episodes. This difference is statistically significant at the 10% level. Given that we do not know exactly when learning starts to improve, it is possible that this is really picking up that episodes with the largest learning improvements start as a new administrative cycle begins.

Figure 4: Change in reading scores by administration: Low- and middle-income countries

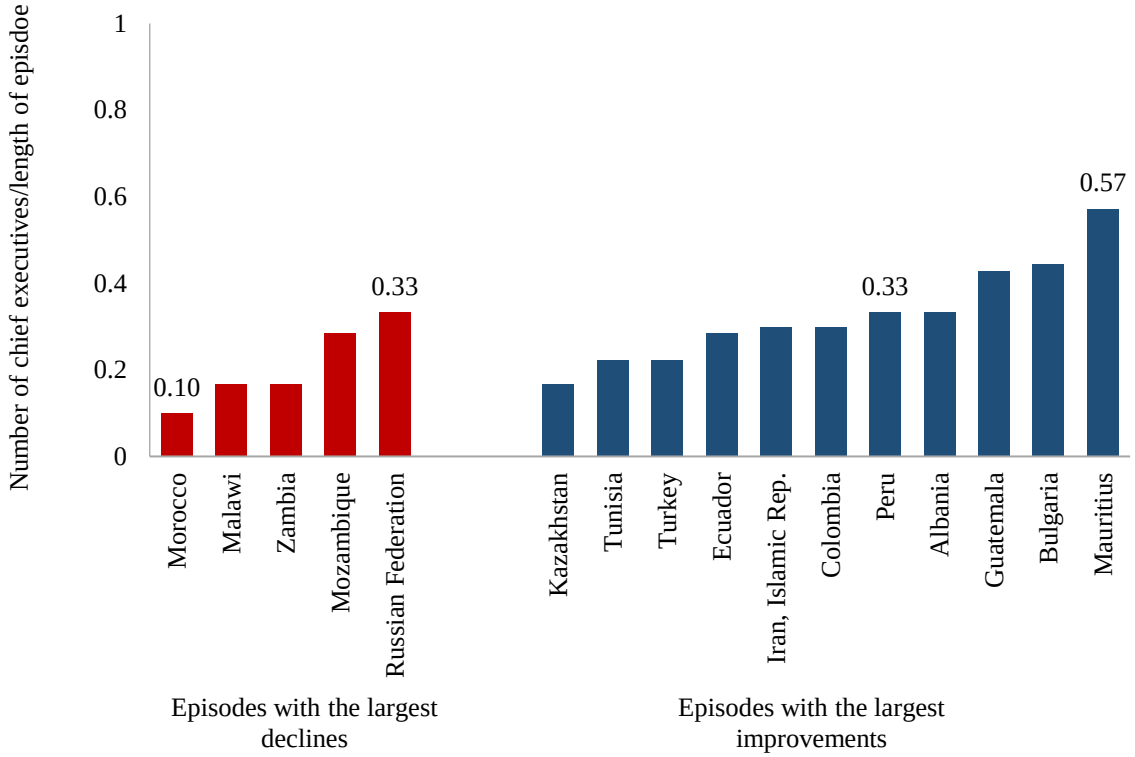


Notes: Bars above the gray line denote top 20% performers.

Another political factor that may help sustain learning improvements is the administration's continuity and stability. Successfully implementing reforms to improve learning takes time, and the longer an administration is in office the more likely they will have the time to fully implement their reform program. While one would expect low administration turnover to provide stability in implementing reform, thereby improving learning outcomes, the data suggests otherwise.

Among low- and middle-income countries, the average number of administrations adjusted for the length of the episode is higher among episodes where learning improved, compared to episodes where learning declined. This difference is statistically significant at the 90% level (see Figure 5). For example, Guatemala had three presidents over a 7-year period of improved learning, while Malawi had only two presidents over a 12-year period and experienced a decline in learning. No such difference exists among episodes in high-income countries (HICs). The finding is consistent with the fact that the best episodes were more likely to begin toward the end of an administration and closer to an election year.

Figure 5: Average number of chief executives per year by episode type, low- and middle-income countries



Notes: Red bars denote the worst episodes, while dark blue bars denote the best episodes with data on the number of different administrations during the episode. Y-axis denotes the number of chief executives over the number of years in the episode.

3.7. Summary

The database analysis suggests some interesting findings between the education, economic, and political context and learning trends:

- While there are many similarities in the education reforms associated with episodes of improving or declining learning in low- and middle-income countries, there are some notable differences. For example, compared to episodes where learning had declined, curriculum, school-based management, learning assessment, and structural reforms were more common in episodes where learning had improved.

- There were also differences in the detail of broad reform efforts. While building institutional capacity was the most commonly introduced reform in low- and middle-income countries,

activities to strengthen financial management systems were more common in episodes where learning improved.

- Initial levels of GDP per capita and economic growth were not significantly different between episodes where learning had improved and episodes where it had declined.

- There was no strong evidence of a trade-off between improvements in learning and improvements in access.

- Spending per student tended to rise more quickly in episodes where learning improved.

- Levels of democracy and the characteristics of a country's political settlement do not seem to have a stronger association with episodes of improved learning compared to episodes where learning outcomes deteriorated.

- While there seems to be some relationship between the timing of reforms in the political cycle and improvements in learning, our information was insufficient to untangle whether successful reforms were introduced at the beginning or the end of a cycle. However, government turnover was not more strongly associated with episodes of learning decline.

The findings in this section are exploratory and are an attempt to look at simple bivariate associations between changes in learning and the education, political, and economic context that was in place at the time. They are in no way causal and exclude many factors that are likely to be important because we were unable to collect cross-country information. However, the results do illustrate some patterns that can be useful in framing claims about the importance of different factors on improved learning. The next section looks more closely at reform episodes in a set of case study countries to provide a richer analysis and assessment of the underlying factors that have driven improvements in learning outcomes.

4. Studying cases of education reform: A synthesis of nine case studies

4.1. Approach

This section qualitatively explores some of the technical and political challenges that states face in their efforts to improve education quality. We offer a thematic synthesis of findings from nine country-level case studies. This synthesis focuses on strategies for deploying information, working with coalitions, and innovating to create opportunities for strategic change. To produce this synthesis, we used the following countries as case studies: Bangladesh, Brazil, Chile, England, Indonesia, Mexico, Peru, South Africa, and Tanzania. Summary characteristics for each country are presented in Table 9 below. The full-length case studies for each country are located in Annex B.

We purposefully selected the nine countries based on the principle of maximum variation in learning trends, income status, and types of reforms.[45] This let us to examine how and why political factors contributed to reforms by drawing on existing data and secondary sources.[46] Two key criteria for case study selection were the availability of learning data over time and whether reforms and reform processes could be identified and be plausibly linked to learning trends. Most country cases we selected came from the database described in Section 3. However, we also added two other countries that were not in the database: Bangladesh, which does not seem to be in any cross-national learning assessments, and England, which offers an example of reforms from a developed country.[47] We then used existing data and secondary sources to produce brief case studies for each country.

The synthesis is organized into three major themes in alignment with the 2018 WDR's focus on systemic change and learning.[48] These themes include (1) information and metrics, (2) coalitions and incentives, and (3) innovation and agility. It then offers a brief discussion, summarizing lessons learned, and considers implications for how the relationship between policy adoption and implementation of education policy can be best understood in a political-economic context.

[45] Patton (2002).
[46] Yin (2009).
[47] In the case of Tanzania, the learning period captured in the database is 2000–2007 (based on SACMEQ participation), but in the case study we explore reforms introduced in 2013, drawing from Early Grade Reading Assessment (EGRA) data between 2012–2016. For Indonesia, the database shows an upward learning trend (based on the PISA reading), but our conclusion in the case study is mixed. We do not see similar improvements in math or Trends in International Mathematics and Science (TIMSS) over the same period. For Indonesia, the database shows an upward learning trend (based on the PISA reading), but our conclusion in the case study is also mixed.
[48] See World Bank (2018, ch. 11).

Table 9: Synthesis country, reforms, political leadership, and economic context

	Bangladesh	Brazil	Chile	England	Indonesia	Mexico	Peru	South Africa	Tanzania
Reforms (Reform period)	PEDP (2000–2014; 2011–2015).	FUNDEB; IDEB (2006–2015; 2000–2012).	SNED; AVDI (1995–2005).	2007 National Literacy Strategy (1995–2015).	2005 Teacher and Lecturer Law (2005 onwards).	Carrera Magisterial; Alliance for Quality of Education; Education Reform Bill (1992–2013).	2007 Teacher's Reform (2000–2015).	Curriculum 2005 (2000–2014).	BRNEd (2012–2016).
Change in leadership during reforms/ Major political incidents around reform period	Awami League has remained in power, with exception of a 2-year period of military rule; Nationalization of private schools strengthened teachers' union's power.	Cardoso government (1996–2003) introduced radical education reforms; progressive administrations of da Silva (2002) and Rousseff (2011).	Three different administrations of the same coalition (1990–2005).	New Labour Party assumed power in 1997.	None: Yudhoyono president from 2004 to 2014.	Five different presidents since 1992. PRI lost elections to PAN in 2000 but regained presidency in 2012; 2013 reforms marked resurgence of PRI presidency after a 2-decades long hiatus.	Toledo (2001–2006); García (2006–2011); Humala (2011 to present); then Candidate García's confrontation with a major teachers' union.	End of apartheid in 1994; shift from Government of National Unity to African National Congress.	Kikwete is president from 2005 to 2015; CCM party wants high-profile education program to win support in 2015.
Learning outcome trends	NSA has shown no improvement since 2011 in either 3rd grade or 5th grade; 25% of 5th graders performed at their grade level in 2013 in Bangla.	Mid-2000–2012: PISA math scores sharply increased 30 points. Both disadvantaged and advantaged students increased performances on PISA test.	2000–2015: increases in PISA reading and math tests. 2000–2006: biggest increase in both subjects occurred, coinciding with introduction of teacher incentive and evaluation reforms.	1995–2015: proportion of students in grade 4 reaching intermediate benchmark in the TIMSS assessment of math increased from 54% to 80%.	2000–2012: PISA reading scores showed steady but small improvements. 2006–2012: scores for math (from PISA and TIMSS) declined.	Aside from a drop in 2003, PISA reading scores improved from 420 in 2000 to 423 in 2015. 2006–2012: PISA math scores improved, but means dipped significantly after 2009.	2000–2012: PISA scores in math increased by two full school years; reading improved by one full school year (327 to 384).	2002–2011: students scoring in highest achievement levels on TIMMS increased; students scoring in lowest achievement levels decreased for ninth graders.	2013–2016: EGRA and EGMA scores improved. Students who achieved the "progressing reader" level increased 40%; those approaching the math benchmark increased 18%.
Trend in enrollment	2005–2011: primary GER increased from 99% to 112%. Secondary NER went from 45% in 2000 to 53% in 2013.	2001–2012: GER decreased from 151% to 130%; NER decreased from 98% to 95%. Net secondary rate was 79% in 2012.	Secondary GER was 87% during the beginning of the learning episode. 2000–2014: secondary enrollment increased to 100%.	Primary education near universal throughout the reform period. In 1999, NER in secondary was 78% and was 77% in 2011 (data for UK).	No change in primary enrollment. Access to lower secondary increased from 71% in 2001 to 89% in 2012. Between 2005 and 2008, it increased from 72% to 82%.	Secondary NER increased since early 2000s. NER (66%) has stagnated since 2007. NER in primary has remained about 95% since 2000. 2000–2014: primary GER decreased by 3% points.	2000–2014: primary NER decreased from 98% to 93%; secondary NER increased from 65% to 78%.	87% NER in primary school in 2000. Secondary GER was 87% in 2000 and 94% in 2014.	In 2012, primary GER was 90%; NER was 84%. Secondary GER was 34%.
Economic growth	2000–2014: annual average GDP growth was 4.3%. 2011–2015: annual average GDP growth was 5%.	2006 and 2015: annual GDP growth averaged 1.6%, down from a 2.0% average rate over 5 years preceding the reform period. 2000–2012: rate was 2.4%; compares favorably with the previous 5 years, when growth averaged only 0.5%.	2000 and 2015: Chile's per-capita GDP grew at an average annual rate of approximately 3%. The only year with negative growth was 2009, but the economy recovered swiftly.	1995: average annual growth rate was 2.5%. After a sharp dip in 2009, 2015 growth rate was 2.3%. (data for UK).	2000–2012: steady increase in real GDP per capita. Average annual growth rate in real GDP per capita was 4%.	2000–2015: average annual growth rate was less than 1%. After a sharp dip in 2009, growth has since averaged around 2.7%.	2000–2015: annual average GDP growth was 4.8%, a strong increase compared to previous 5-year period, when annual GDP growth averaged 1.1%.	2000–2014: annual average GDP growth was 1.8%.	2012–2016: annual average GDP growth was 3.5%, a strong increase compared to the previous 5-year period, when annual GDP growth averaged 2.8%.
Trend in education spending; Spending per student	2006–2009: primary per-student spending increased 14%. 2000–2011: secondary per-student spending doubled.	2000–2012: primary education spending per student increased 62%. 2002–2012: secondary education spending increased 66%.	2000–2013: secondary per-student spending increased nearly 60%. 2000: per-student spending in constant PPP dollars was 2,078. In 2013, it was 3,320.	1998: per-student expenditure of GDP was 12% (primary) and 23% (secondary). In 2015, it was 25% (primary) and 23% (secondary) (data for UK).	A 2003 law set education spending at 20% of national budget. 2000–2006: spending doubled. 2007–2012: per-student spending increased 24% in primary, 34% in secondary.	1999–2011: per-student spending at the secondary level increased 43%. Spending at the primary level increased 35%.	2000–2014: primary spending per student nearly tripled. Secondary education spending per student increased 56%.	2000–2014: primary education spending per student increased 32%. 2000–2014: secondary education spending per student increased 40%.	2008–2012: education was 5% of GDP. 2013/2014: 70% of education budget spending allocated to district-level governments. Per-student spending is not available.

Notes: AVDI stands for Asignacion Variable por Desempeno Individual, BRNEd for Big Results Now Education, CCM for Chama Cha Mapinduzi, EGMA for Early Grade Mathematics Assessment, EGRA for Early Grade Reading Assessment, FUNDEB for Fund for Maintenance and Development of Basic Education and Valuation of Education Professionals, GER for gross enrollment rate, IDEB for Basic Education Development Index, NER for national enrollment rate, NSA for National Student Assessments, PAN for National Action Party, PEDP for Primary Education Development Program, PISA for Program for International Student Assessment, PRI for Institutional Revolutionary Party, SNED for Sistema Nacional de Evaluacion del Desempeno de los Establecimientos Educacionales, and TIMSS forTrends in International Mathematics and Science.

4.2. Information, metrics, and evidence

Information and metrics can guide the development, implementation, and evaluation of policies aimed at improving quality. Information and metrics can include learning outcomes such as some of the international examinations described in Section 3. Metrics can take other forms as well. Key determinants of learning, such as access to pre-primary school, dropout and repetition rates, poverty, and nutrition indicators, are examples of evidence that can contribute to a broader understanding of issues relevant to improving learning.[49]

Key stakeholders can use these forms of evidence to advocate for education quality in different ways. On a technical basis, information and metrics can help with planning, evaluation, and accountability. Decision-makers can better understand policy effectiveness, establish benchmarks, and evaluate performance. We can learn from implementation successes, as key components or principles can be replicated for other programs. We can also learn from failures and challenges to determine how and why this happened and can consider what aspects of an ineffective program should be changed, replaced, or dropped.

Information and learning metrics can also make a political case for improving quality, and the publicity of poor outcomes can lead to calls for action and change. Take the Global Commission for Education Financing and the 2018 WDR, for example. Both were borne out of a global body of evidence suggesting the high enrollment/low quality paradox in basic education has contributed to a global learning crisis that governments and civil society must respond to through action and investment.[50]

Yet, the relationship between information and decision-making is not always straightforward. Metrics are not always harmonized with action. Politics can drive learning-focused reforms, but it can also have the opposite effect. Evidence can be subject to manipulation and opportunism. Metrics can be misleading, manipulated, or used to promote interests in the education sector that are not necessarily aligned with learning or equality. For example, local education officers may receive incentives to improve their jurisdiction's examination performance, but they may do so by not allowing low performers to sit for the exam in the first place.[51]

Despite these challenges, one reason to keep information and learning metrics as the central driver of a quality agenda is to consider what might happen if information did *not* drive decision-making. Failing to place information and learning metrics at the center of planning and accountability can threaten the potency of education quality reforms, opening them up to powerful interest groups or ideologies that may not be aligned with learning. To be sure, information and learning metrics will never offer a perfect guide to policy, but it is crucial to ensure that systemic effectiveness aligns with learning. This section draws on case study material to review the politics of information and learning metrics. It is divided into two subsections, including (1) information and the political incentives to improve learning and (2) the role of information to improve incentives within schools.

[49] World Bank (2018).
[50] International Commission on Financing Global Education Opportunity (2017); World Bank (2018).
[51] Williams (2016); Laterite (2017).

4.2.1. Information and the political incentives to improve learning

Governments have used a variety of approaches to leverage information. Some countries used learning metrics to improve accountability at local levels. In Brazil, the government used performance benchmarks to improve learning outcomes in underperforming schools. Performance targets were established in each state, municipality, and school across the country. Resources were then targeted to low-performing schools, who received support to hire qualified school administrators and to establish a mentorship program that paired top- and low-performing schools. Schools were also publicly ranked so that households could hold their children's schools accountable.

In England, the government used data to drive a sense of bipartisan urgency to improve literacy rates. Before the 1997 education reform that eventually led to better outcomes, the country's Office for Standards in Education (Ofsted) established monitoring and evaluation measures to track learning over time. It was then up to whichever political party that was elected in 1997 to improve children's learning. After the 1997 election, the National Learning Strategy was introduced by the Labour Party. The strategy was informed by a successful literacy project that had been implemented on a smaller scale, the National Learning Project. The project had shown efficacy through case control studies on a subnational level.[52] Throughout the reform period, the government used information and metrics to refine and modify the National Learning Strategy. In addition, the government used data from multiple sources and studies to reduce the possibility of bias.[53]

Effective communication strategies often included using information and learning metrics to build support from key stakeholders, including the public. In Peru, the government used learning outcomes to build public support for a merit-based teacher pay system. The reform was opposed by a teachers' union, which launched a strike. Rather than cede to the union's demands, Peru's president publicized the poor results of sixth-grade teachers in math and reading, pinning the low outcomes on poor teacher quality—and the poor teacher quality on unions. This data-driven approach to advocacy succeeded in swinging public support in favor of quality education reforms. A poll showed that 74% of people were convinced that the government's proposed reform to teacher accountability would improve the public education system.[54]

Establishing a data-driven system does not automatically guarantee that the data will be used to introduce quality reforms, however. The use of information and metrics, like anything else, is subject to politics.[55] It can also be affected by a lack of accountability, the interests of powerful stakeholders, lack of follow-up, and so on. The misuse of information can create perverse incentives for those in the education sector. When planning education quality reforms in Indonesia, information and learning metrics were not prioritized; instead, the needs and interests of teachers were. Poor learning outcomes were evidently not the central consideration driving the reform

[52] Machin and McNally (2004).
[53] Machin and McNally (2008).
[54] Cuenca (2017).
[55] Sandefur and Glassman (2015).

process. Rather, national and local politics drove public perception to focus on the welfare of teachers rather than students' learning needs. The need to improve quality was not communicated to other actors, such as parents and community members, who may have supported the introduction of competency tests. This underuse of information and learning metrics was a missed chance to build a long-lasting coalition that centered upon raising learning standards.

4.2.2. The role of information to improve the incentives within schools

Governments can use data to improve school-level operations and accountability and to ensure that resources are targeted to locally relevant issues. In Tanzania, the public expressed collective concern about the declining pass rates on national exams. Results on the country's EGRA and Early Grade Mathematics Assessment (EGMA) also highlighted that low education quality was a systemic issue. National government officials shared this concern, particularly due to an upcoming national election in 2015. The government extensively used data to introduce an education reform, setting benchmarks to monitor progress, rank schools, and channel resources based upon performance. The government also published an online interactive district-level map that showed examination performance as a way to improve public accountability.[56]

Governments can use information to link teacher performance with salary increases. Galvanized by low learning outcomes, for example, leaders in Chile introduced a series of reforms that included programs holding schools and educators accountable. The government used data to evaluate schools based upon five variables: performance in the national student assessment; completion and equality of opportunity; integration of teachers and parents; improvements in staffing conditions; and initiatives on pedagogical activities, school development plans, and the creation of teacher councils.

Chilean leaders also introduced a teacher evaluation system in 2004. Teachers were evaluated every four years based on a competency examination as well as qualitative assessments consisting of classroom practice, self-assessments, and peer evaluations. Teachers were categorized based upon performance, with the top two performance categories rewarded with the possibility of promotion. Top-performing teachers also received individual bonuses. And to make the reform palatable to teachers' unions, data-driven forms of local accountability were only gradually introduced. In 1996 the Chilean Ministry of Education introduced a program to link bonus pay for teachers to school-based performance. Even a decade later, teachers still had the chance to opt into evaluation in exchange for bonuses and individual incentives.

4.3. Coalitions and incentives

The introduction and implementation of reforms are also driven by, and contingent on, the buy-in of powerful stakeholders. Even when evidence and learning metrics are available, there is no guarantee that this evidence will be acted upon. Decision-makers and other key stakeholder do not always align their interests with a quality agenda. They have formal and informal coalitions who

[56] World Bank (2014).

often have other priorities and incentives that they need to consider to maintain power, keep others satisfied, and win elections.[57]

Yet, mobilizing powerful stakeholders, both individuals and groups, is essential to policy reform. Hickey and colleagues[58] note that the elite are "often more adept at attracting public goods provision and at maintaining a better quality of service delivery through social accountability mechanisms." The incentives and ideas of these key stakeholders can reveal what kind of reforms they introduce and how well they finance, implement, and assess these reforms. Coalitions can take many forms. Consultations are about the actual substance of the decisions made, but they are also about working to ensure that groups are understood and heard, their interests have been accounted for, and that the decisions will contribute to change rather than be an impediment to it. In this section we turn to examine (1) the role of coalitions in improving education quality and learning outcomes and (2) the role that incentives and individuals have in advancing a quality agenda.

4.3.1. Coalitions and their role in improving education quality and learning outcomes

The role and form of coalitions were diverse across the case studies. The countries that most effectively introduced and sustained reforms aimed to be inclusive of various groups and stakeholders at different levels of government and civil society. Those that failed to do so faced difficulties when it came to implementing reform because buy-in was not widely shared.

Across most of the country cases, one of the most important stakeholders was teachers' unions. Teachers' unions often had a great deal of power and influence, but their interests were not always aligned with efforts to improve student learning. Their impact typically followed from their position as a large decentralized voting group, whose members played one of the most important roles in delivering better quality education to children. Teachers in government systems are often state employees working in their own communities. Sometimes they provided other important functions for local governments such as working at polling stations. Thus, if a politician or political party wished to stay in power, they needed the support of teachers' unions.

In Bangladesh, for example, there were over a dozen teachers' unions consisting of about half a million members.[59] The teachers' unions were successful in getting the government to nationalize all registered nongovernment primary schools. This meant that all teacher salaries, ranks, and status became part of the government systems. Teachers now comprise nearly half a million middle-class voters, "giving them considerable clout in the national policy space"[60] while also guarding against accountability based upon learning outcomes.

Alienating unions ran the risk of undermining an electoral base whose support politicians would need. But ceding to their demands could have meant failing to introduce policies aimed at incentivizing quality improvements. As discussed above, during his reelection campaign, Peru's

[57] Kosack (2012).
[58] Hickey, Sen, and Bukenya (2015, p. 6).
[59] Hossainet al. (2017, p. 11); Richards and Vining (2016).
[60] Hossain et al. (2017, p. 16).

president publicly blamed teachers' unions for the country's low learning outcomes. He was successful at convincing the public to side with him to support reforms, which included introducing teacher competency exams.[61] But by taking a confrontational stance to fight unions, the reform process alienated and excluded a vital stakeholder: the teachers who would ultimately implement the reform.

Similarly, in Mexico, the government struggled to find an effective way work with the teachers' union, which had curtailed the effectiveness of its efforts to introduce reforms to improve quality. The country introduced major reforms to address accountability and to improve learning outcomes. However, there was a recurrent feature across the reform episodes: uneven power dynamics between the teachers' unions and education authorities. For example, in 1992 the central government, state governors, and teachers' union signed an agreement to decentralize system operations and introduce a merit-based pay program.

In the early 2000s a new administration took power, and they added another agreement to improve education quality. And in 2008 they created another alliance with teachers' unions to amend teacher policies. However, misalignments in reform design and failure to address political constraints hampered implementation. In 2013, the administration adopted a more confrontational approach and introduced a reform package to overhaul the teacher hiring and evaluation system, one that aimed to reduce patronage and factionalism while improving accountability. But in the absence of broad-based consultations and effective communication, these reforms were met with stiff resistance and have not been fully implemented.

In addition to teachers' unions, there were other types of elite coalitions that impacted national efforts to improve quality. In South Africa, for example, the new post-apartheid government, African National Congress (ANC), had to contend with the ongoing presence of pre-apartheid government members, the Government of National Unity (GNU). The latter group had more policy making experience, making them a powerful coalition. Even though apartheid had officially ended, GNU members promoted interests that served their their white South African constituents.

One strategy to improve the effectiveness of reforms was to enlist key coalitions during the reform-planning process. To improve learning outcomes in Tanzania, the government convened a consultation committee consisting of some of the most important stakeholders in the country.[62] The committee consisted of 34 members belonging to 31 groups and included important governmental and nongovernmental actors. The teachers' union was one group that had been reluctant to participate in reform efforts. Teachers viewed the proposed training and skills development as opportunities for professional development, but they were also concerned about adding more responsibilities without any significant monetary benefit. To succeed in getting the union on board with the reform, the government acquiesced to the union's wishes for offering monetary and nonmonetary rewards as part of this new reform. As a result, the education reforms were successfully developed and implemented, bolstered by the support of a broad swath of key coalitions, including the union.

[61] Cuenca (2017).
[62] The reform process was modeled after a similar strategy used in Malaysia.

4.3.2. The role of incentives, individuals, and quality education policies

To overcome political obstacles and bring about technical reforms, the buy-in of political elites was essential. The role of specific individuals and incentives could also be influential. The backing of a charismatic leader, for instance, could exert influence to leverage political capital in order to gain public support.

Drawing again on the example of policy reform in Tanzania, the government and the president were incentivized to improve quality for two interrelated reasons: it was an issue that the public cared about, and there was an upcoming election. The introduction of the education reform, called Big Results Now Education (BRNEd), aimed to address the country's low pass rates on the primary and secondary school national examinations.

Under President Jakaya Kikwete, the government approached the problem by using a multi-faceted program aimed to improve the skills of students, teachers, and administrators; making education fiscal management more efficient; and setting benchmarks to measure learning achievement through strong accountability mechanisms. The president and his party were also positioning itself for the 2015 elections, and introducing a successful reform was seen as a way to win support across the country. Regional and district commissioners responsible for overseeing the implementation of BRNEd were mostly members of the ruling political party. Thus, their interests and incentives were aligned with the successful implementation of BRNEd.

Brazil's government used a combination of coalition tactics and elite leadership to push through an education reform program that focused on improving quality through a decentralized system. Several important political figures leveraged their political capital to introduce decentralized quality reform. The president and first lady used selective coalition building to get congressional approval. At the time, state politicians had little incentive to support a decentralized education system, so the elites took advantage to get this approval. Their strategy hinged upon introducing strong incentives for local authorities to endorse primary education. The reform was approved by ministerial decree rather than congressional legislation. It transferred power to states and municipalities, and coalitions were forged with local counterparts at three levels: states, municipalities, and school units.

Innovation and partnership building spurred the reform process in Brazil. The accountability made through decentralization mechanisms and learning targets have been sustained. The swift coalition strategies that the ministry had exerted to approve the fund helped manifest accountability and social controls locally; this then helped the decentralization of education provision and helped expand its reach. The result was that local municipalities became electorally accountable for the municipal school system's performance.

Education reforms in Chile were driven in part by the leadership and vision of President Colegio. However, they were also informed by the technical knowledge and leadership of well-qualified leaders in the education sector. Four out of the seven education ministers from 1990 to 2005 held doctorate degrees in economics or education, and one of the ministers had experience working with the Colegio government to solve internal financial issues, which helped maintain good

working relations. Solid technical leadership in the ministries of education and finance allowed the two ministries to collaborate and design sound policies.

4.4. Innovation and agility

Education systems should be designed to improve quality and equity. Built into this process is the need for systems to be responsive and flexible enough to adapt to shifting political, economic, or environmental conditions. In this section we turn to explore the themes of innovation and agility. Being innovative and agile means being able to change in response to different circumstances and opportunities. These can range dramatically such as being able to adapt to innovations or build off well-performing aspects of the sector to find cost-cutting measures by strengthening efficiency. As the 2018 World Development Report notes, "Exploring the well-performing parts of any education system can reveal technically and politically feasible approaches to the problems systems face in improving learning."[63]

The principle of being structured yet flexible is to know that while systems should be driven by evidence, they are embedded in a broader political and economic context that is constantly in flux. In other words, shorter-term evidence-based planning must align with longer-term goals of improved learning, and yet the shorter-term strategy to achieve these goals might need to change, should circumstances permit. A major political change (e.g., end of apartheid), an economic downturn or upswing, or a teachers' union unexpected change in leadership can each be events that education reforms must respond to in order to maintain a focus on improved learning. The WDR characterized systems that can innovate and respond as "open," noting that such systems "pay more attention to overall outcomes and reward progress in raising outcomes are more likely to see greater innovation and the diffusion of new approaches across the education system."

We explore these issues across three sections including (1) identifying solutions to local problems, (2) innovative and adaptive policies, and (3) finding critical moments in the development of education systems to improve quality.

4.4.1. Strategies for identifying solutions to local problems

Ongoing dialog between national and local levels of governments was often essential to bring about sustained change and accountability processes. Developing local approaches to assessment and quality improvement needed to reflect local priorities. In Tanzania, for instance, communities weren't explicitly concerned with international assessment scores. Rather, what was more meaningful for parents was that their child did well on the national examination and attend a well-ranked school. After the government took this into account, and the education reforms then took these local priorities into account, they added incentives and accountability measures that allowed examination scores to function as a proxy for quality.

Brazil's experience of decentralizing its education system offers another example for how national-level reforms can be designed to respond to local needs—and it serves as a lesson that it

[63] World Bank (2018, p. 25).

takes time for innovative policy reforms to mature and take root. Brazil's most successful reforms occurred over two periods, with strong central leadership roles spearheading the process. The Cardoso government (1996–2003) introduced radical education reforms, and the Lula da Silva (2002) and Dilma Rousseff (2011) administrations followed suit, building upon prior reforms to expand coverage and improve quality.

Brazil's second reform wave consolidated and strengthened access and quality. Coalitions and partnerships helped lead reform incentives, along with a sustained commitment to continuing the general trajectory of education quality. The government developed a policy to respond to local needs based on decentralization, and it was iterative in the sense that local municipalities were empowered (and held accountable) for improving quality and attendance.

As a result, municipal decentralization redistributed federal resources according to need. Subsequent policies expanded student coverage,[64] strengthened local accountability, and introduced assessments and monitoring tools. Through this process, Brazil went from being one of the lowest performers in the PISA to experiencing sustained improvement.

4.4.2. Innovative and adaptive policy making and implementation

Across the case studies, one characteristic of the most successful reforms was that they were adaptable and amenable to modification and change. The reforms may have been bolstered by a charismatic leader or a change in the political settlement, but they were most effective if governments sustained their commitment (and general trajectory) toward improving education quality. Allowing for some iteration and flexibility permitted polices (and policy makers) to not "get it right" the first time around. Rather, they oriented the trajectory of the system in a way that could adapt and respond to systems, actors, and information that were in a state of flux.

In the case of England, its commitment to improving learning outcomes was sustained across political settlements (i.e., Conservative and New Labour), but its system was also flexible enough to adapt to existing political and ideological conditions. During the 1997 elections, the issue of poor education outcomes was a concern for both major political parties. The National Literacy Project (precursor to the National Literacy Strategy) had actually been first introduced under the Conservative government.[65]

Before the 1997 election, an outline of the National Literacy Strategy was launched, and the government shared a summary of the report with each primary school across the country.[66] There was sustained pressure to improve children's learning, and both parties recognized the need for reforms. In the early 1990s, the government had already introduced nonpartisan monitoring and accountability measures to track learning.[67] In 1998, the government published a national strategy, one that they felt would be technically and politically viable, delineating its plan to improve learning and quality for all.[68]

[64] The major reform was the Primary Education Fund (FUNDEF, later known as FUNDEB).
[65] Machin and McNally (2008).
[66] Stannard and Huxford (2007).
[67] Machin and McNally (2008).
[68] The numeracy strategy was introduced in 1999.

Reforms that link teacher performance to teacher pay are typically short lived and have a limited track record of success. However, Chile's experience in negotiating and implementing teacher performance incentive reforms during the mid-1990s to early 2000s stands as an outlier. Despite an active teachers' union, the Ministry of Education managed to pass several reforms directly related to teacher performance during this period. The reform process highlighted the benefits of using a phased approach to introduce a technically sound and innovative program.

Chile implemented school-based bonus pay and individual pay incentive programs tied to performance assessments that are still in practice. Although the union at first resisted proposals for introducing collective performance incentives,[69] they eventually supported the program and helped design later policies on individual-level evaluations. By late 2000s, about 15%–25% of teacher salary was linked to group- and individual-level performance-based incentives.[70] This approach of sequencing teacher incentives coincided with improved learning outcomes. It is not surprising that these reforms remained relatively stable over time.

4.4.3. Critical moments in developing education systems to improve quality

Studying the introduction, implementation, and effectiveness of education reforms needed to be understood in political-economic and historical context. In doing so, several case studies revealed "critical moments" in the country during which time reforms were introduced, and these critical moments varied considerably. The obvious one could be a major change in the political settlement accompanied by a commitment to reform. But critical moments could also take other forms as well, such as a report highlighting low learning outcomes or a moment of weakness (or empathy) in the capacity of a teachers' union related to introducing teacher accountability measures to improve learning.

Perhaps one of the most clear-cut critical moments occurred in South Africa. The end of apartheid brought a period of immense social, political, and economic transition for the country and education system. The ANC championed racial equality, including the introduction of an integrated and equitable national education system. For example, in 1996, the country's first post-apartheid Minister of Education introduced the National Education Policy Act which established a national education system to be used by the whole country.[71] The act also decentralized the national system and gave provinces the autonomy to deliver education and training in ways they needed. The ANC came into power at a time when South Africa was deeply divided along racial lines and class inequalities. The new government was eager to develop the basic education system in a way that could symbolize its commitment to all South Africans.

In England, the political milieu presented a perfect storm of evidence, opportunity, and public support to introduce the policy reform. There was already demand from the public (and political will) to improve learning outcomes. New Labour came into power, in part, on their pledge to improve education quality. While they had the political capital to introduce major quality reform, they also had the luxury of not starting from scratch. Instead, they took advantage of the critical

[69] Mizala and Schneider (2014).
[70] Mizala and Schneider (2014).
[71] OECD (2008).

moment: they built on an existing program (National Learning Project), one with demonstrated efficacy,[72] to introduce a transformative education reform to improve literacy.

4.5. Summary

Findings from the case studies illustrated some of the complexities that shaped the ability of governments to deliver quality reforms for children, particularly in relation to data and information, coalitions and incentives, and forms of innovation and agility. Lessons learned include the following:

- Elite buy-in was essential for introducing technical reforms. Successful reforms were not contingent on one charismatic leader. Rather, quality reforms needed to be introduced by leaders who were informed by well-qualified experts.
- Governments committed to improving learning outcomes wanted to make information drive their decision-making. Information and metrics helped build demand for reforms that could track learning, improve the curriculum, assess performance, and so on.
- Information could also drive communication strategies and be used to make the argument for reforms and gain the support of key actors, including the public. An effective communication strategy also counters the spread of misinformation by groups who may oppose the reform.
- Effective coalitions enlisted the buy-in of stakeholders that would be vital for introducing and implementing the reform. Teachers' unions were often a powerful opposition group. Taking a confrontational stance to fight unions ran the risk of alienating the group of teachers who must implement the reform. In the country cases we looked at, teachers' unions often succeeded in blocking efforts to introduce quality reforms.
- The most effective reforms were focused but flexible. Policies were effective if they had a clear direction but could also be changed or re-envisioned as external conditions changed. Reforms can build on one another over a longer period of time, gradually offering greater levels of sophistication to the system in a way that can incrementally improve learning.
- "Critical moments" could give rise to quality-focused reforms. A change in leadership, a major political shift in the country (e.g., end of apartheid), or a report highlighting low learning outcomes can compel leaders to accept reforms.
- Sequencing popular reforms with those that are more likely to be opposed helps to increase acceptance of reforms that were less popular. In the case of teacher accountability, starting with collective incentives can help build an evaluation culture that can make it easier to begin individual incentives later on. Such an approach was more likely to guarantee implementation of meaningful reforms rather than launching a drastic policy change that may be rejected at the outset.

[72] Machin and McNally (2008).

5. Discussion and conclusion

In this study, we attempted to study the systemic factors that contributed to improved learning or to declines. To be sure, our study findings must be read as exploratory. We do not offer definitive conclusions about the relationship between the different variables and countries from the dataset or the case study synthesis. We note, for example, that the dataset was far from perfect. We had to rely on the information available to us (e.g., political settlement classifications based off of data from 2000). However, by offering some simple associations and hypothesizing the relationships of variables, we hope this study can offer a helpful perspective that can complement other recent attempts to study the politics of education quality reforms.[73]

We also hope that the questions generated from this work can lead to more granular work to explore these associations in greater levels of detail. Similarly, the case study synthesis was an attempt to locate our investigation of policy reforms in their real-life contexts, but the conclusions from this section must be interpreted with modesty given their brevity and reliance on secondary sources. What they can do, though, is help locate key themes and associations within the complex political and historical context in which they exist.

There are other ways of studying education policy reform that would benefit from further investigation. For example, future research can study other aspects of quality such as repetition and dropout rates and teacher qualifications. Given the broad reach of the study, it was difficult for us to fully attend to some issues that may have more salience in some countries than others. We could not capture the significance of conflict, disaster, or other factors that may contribute to political instability or tell the types of reforms that were (or were not) introduced as a result. Similarly, the relative brevity of the case studies limited the depth in which they could engage with the enduring effects of colonialism or structural violence.[74]

Further research could also explore the rise of (and demand for) low-fee private schools and whether and how this impacts the drivers and demand for quality public basic education. Finally, an analysis of politics and reform may benefit from more cross-sectoral investigation. Such a study could look at whether trends in education reform align with those found in public health or social protection. Knowing about whether the political drivers of delivering quality education reforms align with the incentives in other arenas would be instructive for explaining outcomes seen within and across countries.

Type of education reforms

Our findings from the database on whether specific types of education reforms predict learning trends indicate that the type of education reforms per se may not determine learning outcomes. We find that reforms related to certifying teachers, strengthening financial management and

[73] Recent works of note include a series of working papers published by Research on Improving Systems of Education (RISE) (www.riseprogramme.org) and the Effective States and Inclusive Development (ESID) program at the University of Manchester (http://www.effective-states.org/?s=education); an edited volume on governance and basic education in South Africa by Levy et al. (2018); an edited volume by Hickey and Hossain (2019) that includes country case studies of Bangladesh, Cambodia, Ghana Rwanda, South Africa, and Uganda; and a review of the politics of quality reforms related to sustainable development goals by Bruns, Macdonald, and Schneider (2019).

[74] Sriprakash, Tikly, and Walker (2019).

accountability, and school-based management (SBM) reforms were more likely to coincide with episodes of improvement. But aside from these differences, assessing the relative rank of different reforms across episode types does not suggest that countries favored one set of reforms during periods of learning improvement versus in periods of decline. This suggests that learning trends also depend on whether reform components are technically or politically aligned toward improving learning or are coherent with existing programs and policies.

On the other hand, one reason why countries may have favored similar types of reforms during episodes of learning improvement, as well as declines, is that it is politically appealing to continue certain activities such as those related to improving access for all. A shortcoming in our dataset is that there is no distinction between whether the reforms were implemented or were just planned during our period of interest.

Economic conditions

While there are reasons to believe that the right economic conditions can contribute toward improving education quality, we do not find a strong correlation between economic growth and learning episode type. Moreover, the absence of a significant correlation between economic growth and learning trends, but a potential relationship between the change in growth rate and learning, could suggest that exogenous factors affecting changes in economic conditions beyond economic growth itself may matter more for improving learning.

Potential factors could include changes in political and administrative conditions that may directly or indirectly affect the education system. For example, Hausmann, Pritchett, and Rodrik find that political regime change is a strong predictor of growth accelerations.[75] Among low- and middle-income countries, we do find that the best episodes are likely to start in wealthier countries, but this is not true among HICs. These findings, or lack thereof, further suggest that while countries with a higher GDP may have conditions to foster improvement in learning outcomes, the potential benefits from increased economic growth for education quality may not always be realized.

Spending on education

The lack of a consistent relationship between education spending and learning outcomes could reflect potential misalignments in education systems in countries where spending increased but commensurate improvements in learning did not happen. For example, countries may have invested in expanding secondary level schools without taking steps to increase the number of teachers qualified to teach secondary level courses. Even if spending increases are aligned toward learning, bottlenecks in the flow of funds from the center to the frontlines, or mismatches in fund use, can explain why spending increases do not always result in learning improvements.

Expansion in access

In the case of enrollment in primary and secondary levels, episodes of improvement experienced a lower average increase in enrollment. However, this does not hold when comparing the best and worst learning episodes based on the PISA tests. Although it is difficult to explain the opposite pattern in changes in enrollment and learning trends for the PISA-based episodes, the greater expansion in access during the best episodes could reflect the fact that more countries with the

[75] Hausmann, Pritchett, and Rodrik (2005).

worst episodes included HICs that already had a higher secondary gross enrollment rate (GER) at baseline.

Although the small number of countries in each of these groups limit our ability to make definitive conclusions, the trend during episodes based on primary level assessments is consistent with the idea that a greater focus on enrollment may compromise quality improvements, hence damaging learning outcomes.[76] Moreover, a further look at enrollment rates in the individual episodes reveal that all six countries with the best episodes had already achieved universal primary enrollment when the episodes began. This was the case for only two out of the five countries with the worst learning episodes. For example, primary GER in Mozambique increased by 30 pp over seven years from a base of 74% in 2000, and over the same period it also experienced a drop in SACMEQ reading scores by 40 points. Thus, countries that have already achieved universal enrollment may be less likely to prolong focus on access and be more likely to be pressured to focus on quality.

While our findings on the trends in enrollment at the primary and secondary levels by episode type are not consistent, the data suggest that investing in early childhood education (ECE) may be crucial to improving learning outcomes. This finding is consistent with the idea that skills acquired in the early formative stages improve one's capability to learn faster at the primary or secondary levels.[77] The challenge for policy makers in introducing reforms to expand quality ECE coverage is that the learning benefits are not immediately realized. These benefits may also only be visible around seven to ten years later, when pre-primary beneficiaries enter primary schools.

Political conditions

Our finding that the level of democracy is not strongly correlated with trends in learning is not particularly surprising. Dahlum and Knutsen argue that there is indeed little reason to believe why democracies would have incentives to offer higher quality education versus providing more education.[78] Quality improvements are less visible and hence more difficult to attribute to the government's efforts, compared with expansion in access, so the average voter may be unlikely to elect leaders based on quality issues.

Furthermore, if democratically elected governments choose to focus on access, quality may suffer. For example, Stasavage finds that after the Ugandan government abolished school fees in 1996, enrollment increased, but quality dropped from the abrupt shortage in inputs.[79] Moreover, one could posit that resistance to unpopular, but potentially effective education reforms, is more likely in democracies that allow freedom of speech and offer legal or other platforms to mobilize opposition.

On the other hand, we find some surprising evidence: the best learning episodes are more likely to take off closer to an election and near the end of an administration cycle. This may occur due to two different reasons. Leaders may be willing to invest in quality improvements to get reelected, which would go against the idea that democratic leaders do not have incentives to improve quality (e.g., Tanzania). Another reason is that if leaders are nearing the end of their administration without

[76] Stasavage (2005).
[77] Cunha and Heckman (2007).
[78] Dahlum and Knutsen (2017).
[79] Stasavage (2005).

prospect of reelection, they may feel more emboldened to challenge the opposition in introducing difficult reforms.

We also find that among HICs, declines in learning began at a time of a more unified opposition, while the best learning episodes started when opposition was more fractured. This result is plausible since it may be politically easier to implement reforms and policies to improve quality if members of the opposition do not present a unifiedfront. The case studies offer further lessons on the strategies countries have deployed to introduce difficult reforms as well as the conditions that have hindered reform implementation and consistent learning improvements.

Ensuring reforms are well suited for the political environment is important for implementing meaningful reforms. Merit-based pay can be a tool to encourage better performance. But teachers' unions can influence all levels of administration, creating an unfavorable environment for effective implementation. Accounting for such risks before introducing a performance incentive program is important and essential.

Even when reform components may have been technically or politically aligned, an effective communication strategy was also essential. It was important for the state to take control of the policy reform's core message: it intended to improve children's learning. Highlighting reform elements that were perceived to be positive could be crucial to receive buy-in and support and to avoid negative press. Arguments for reform can also draw from learning metrics, which can help focus arguments more on the children's welfare than politics or ideology.

Attempts at reforms that seemed to take quality more seriously based their rationale on learning metrics. Some governments, such as Brazil and England, became a "client" to learning outcomes, allowing information to drive decision-making. Falling pass rates in examinations can be alarming for governments, signifying the deteriorating state of education system in the country. Developing and using learning information appropriately can help build demand for quality reforms as well as allow for well-designed reforms. Attention can be placed on developing effective learning information systems to help make the case for quality-enhancing reforms. Information can also be used to help encourage better school systems, and it can drive new targets spurred on by more engaged citizens. Participating in international assessments benchmarks of accountability were important in many cases, but so too was "local" metrics, such as examination reforms, literacy and numeracy rates, or school rankings.

The mere existence of learning outcomes does not guarantee that they will be used to drive policy. Whether or not learning metrics are used may be best understood in the context of the political settlement. Learning outcomes may be only one factor elites consider when developing policy. Whether or not learning outcomes are actually used requires a broader reading of the coalitions and incentives that shape education policy priorities and to what extent gains in learning metrics feature in this matrix. For instance, central education systems may be adept at collecting information, but whether and how this information is used to improve school administration or pedagogy can vary, particularly in large decentralized systems. Informal political alliances and bargaining can shape the recruitment and hiring of teachers more than information can.

Windows of opportunity for introducing quality reform can be exploited. Throughout the cases, there were several moments in which inflexible systems temporarily relaxed, presenting the government the chance to introduce quality reforms. In South Africa, for example, the ANC took power as the country was transitioning away from apartheid and introduced substantial reforms. In England, the New Labour came into power with the political capital it needed to introduce a massive literacy campaign in schools across the country. In other countries, teachers' unions may experience moments of vulnerability or have a leader who is sympathetic to the needs of learners, presenting a period where new reforms can be introduced.

Teachers' unions and other powerful coalitions must be part of the policy making process if reforms are to be effective and sustained. Teachers' unions often wielded a great deal of power and influence, and their interests typically aligned with teacher welfare rather than learning metrics. Their power flowed largely from their role as a large and decentralized voting group whose members (i.e., teachers) were vital for the success of any reform efforts. Union members were often government employees who sometimes performed other government functions, such as working at polling stations. For politicians to alienate unions ran the risk of undermining an electoral base whose support they depended on. But to cede to their demands could mean failing introduced policies aimed at incentivizing teachers to improve quality.

Teachers' unions often succeeded in obstructing government reforms in the country cases we studied. However, two strategies for working with unions held promise. First, unions needed to be recognized and be engaged with as a key stakeholder early on in developing policy, allowing them to have the chance to buy-in to the reform from the beginning. A second strategy that held promise was phase-in, opt-in approaches to improving incentives and accountability. Opting for incremental changes to policy over time, rather than a drastic change, was a more prudent way to gain acceptance and pass reforms.

The most effective policies were focused but also flexible. Policies were effective if they had a clear direction and could be modified, adapted, and even re-envisioned over time. In Brazil, for example, successive governments introduced different components of the quality education reform such as decentralization, strengthening the capacity of local education systems, or building capacity to monitoring learning metrics. The reforms were not contingent on one charismatic leader (though this helped to jumpstart the process). Rather, the reforms built on one another over a longer period of time, gradually offering greater levels of sophistication to the system in a way that slowly improved learning outcomes.

Teacher accountability measures must be introduced strategically. Reforms to improve learning can often be multi-dimensional. Sequencing reforms with those that are less likely to be opposed helps increase acceptance of reforms that are less popular. In the case of teacher accountability, starting with collective incentives can help build an evaluation culture, making it easier to accomplish individual incentives that could be directly tied to learning. Such an approach is more likely to guarantee the implementation of meaningful reforms rather than launching a drastic policy change that may be rejected at the outset.

A consultative approach to designing reforms, involving various actors in education and across administrative sectors, is likely to build support at time of implementation. The cases of Chile and

Tanzania highlighted the benefits of this type of approach. Consultations can help build consensus and buy-in to the reform efforts, ward off would-be detractors (e.g., teachers' unions), and help in a smoother rollout of the policy.

Elite commitment is necessary. Commitment from higher offices can galvanize momentum for quality reforms. As we saw in Tanzania, for example, the president cared about improving quality at least partly because it was an issue of public interest in light of the upcoming elections. Introducing a successful reform was a way to win support across the country, showing evidence of positive impact of the leadership or the president and his party. Similarly, in England, New Labour came into office determined to show it was serious about improving literacy rates, which had become a national issue. And in Brazil, pivotal political actors pushed the reform agenda ahead. Vision and leadership were major success factors in improving the education system. Hence, strategic and broad-based consultations is necessary to produce comprehensive reform strategies.

One of the most notable findings of this study was the absence of strong, significant relationships between most indicators and long-term learning improvements. But what we can do, more or less, is confirm the absence of a magic bullet when it comes to quality reforms. There probably is not any one component in the education and broader economic system that will lead to sustained changes in learning. All parts of the system must be coherent and aligned toward learning.

6. References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. 2014. *Democracy Does Cause Growth.* Working paper no. 20004. National Bureau of Economic Research.

Andrews, Christina W., and Michiel S. De Vries. 2012. "Poverty and Municipalization of Education: The Analysis of Ideb Results (2005–2009)." *Cadernos de Pesquisa* 42(147): 826–847.

Ansell, Ben W. 2010. *From the Ballot to the Blackboard: The Redistributive Political Economy of Education.* Cambridge: Cambridge University Press.

Avalos, Beatrice, and Jenny Assael. 2006. "Moving from Resistance to Agreement: The Case of the Chilean Teacher Performance Evaluation." *International Journal of Educational Research* 45(4–5): 254–266.

Barber, Michael. 1996. *The Learning Game*. London: Gollancz.

Bils, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review* 90(5): 1160–1183.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2013. *Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education.* Working paper no. 321. Washington, DC: Center for Global Development.

Boli, John, Francisco O. Ramirez, and John W. Meyer. 1985. "Explaining the Origins and Expansion of Mass Education." *Comparative Education Review* 29(2): 145–170.

Bouguen, Adrien, Deon Filmer, Karen Macours, and Sophie Naudeau. 2013. *Impact Evaluation of Three Types of Early Childhood Development Interventions in Cambodia.* Washington, DC: World Bank.

Bray, Mark, and Chad Lykins. 2012. *Shadow Education: Private Supplementary Tutoring and Its Implications for Policy Makers in Asia.* Asian Development Bank.

Brazil Ministry of Education. 2015. *Education for All in Brazil, 2000–2015.* Incheon: UNESCO.

Bruns, Barbara, David Evans, and Javier Luque. 2011. *Achieving World-Class Education in Brazil: The Next Agenda*. Washington, DC: World Bank.

Bruns, Barbara, and Javier Luque. 2014. *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean.* Washington, DC: World Bank Publications.

Bruns, Barbara, Isabel H. Macdonald, and Ben R. Schneider. 2019. "The Politics of Quality Reforms and the Challenges for SDGs in Education." *World Development* 118: 27–38.

Bruns, Barbara, and Ben R. Schneider. 2016. *Managing the Politics of Quality Reforms in Education: Policy Lessons from Global Experience.* International Commission on Financing Global Education Opportunity.

Cameron, Stuart. 2010. *Access to and Exclusion from Primary Education in Slums of Dhaka, Bangladesh.* Brighton: Consortium for Research on Educational Access, Transitions and Equity (CREATE).

CAMPE. 2009. *State of Primary Education in Bangladesh.* Dhaka: Campaign for Popular Education.

———. 2015. *State of Primary Education in Bangladesh.* Dhaka: Campaign for Popular Education.

Canedo, Ana. 2016. "Mexico's Education Reform: What Went Wrong." *Georgetown Public Policy Review*.

Carnoy, Martin, Tatiana Khavenson, Izabel Fonseca, Leandro Costa, and Luana Marotta. 2015. "Is Brazilian Education Improving? Evidence from Pisa and Saeb." *Cadernos de Pesquisa* 45(157): 450–485.

Cassen, Robert, Sandra McNally, and Anna Vignoles. 2015. *Making a Difference in Education: What the Evidence Says.* London: Routledge.

Chang, Mae Chu, Sheldon Shaeffer, Samer Al-Samarrai, Andrew B. Ragatz, Joppe de Ree, and Ritchie Stevenson. 2014. *Teacher Reform in Indonesia: The Role of Politics and Evidence in Policy Making.* Washington, DC: World Bank Group.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2011. *The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood.* Working paper no. 17699. National Bureau of Economic Research.

Chinapah, Vinayagum. 2003. *Monitoring Learning Achievement (MLA) Project in Africa*. Association for the Development of Education in Africa (ADEA).

Chisholm, Linda. 2005a. "The Making of South Africa's National Curriculum Statement." *Journal of Curriculum Studies* 37(2): 193–208.

———. 2005b. "The Politics of Curriculum Review and Revision in South Africa in Regional Context." *Compare: A Journal of Comparative and International Education* 35(1): 79–100.

Chisholm, Linda, and Ramon Leyendecker. 2008. "Curriculum Reform in Post-1990s sub-Saharan Africa." *International Journal of Educational Development* 28(2): 195–205.

Contreras, Dante, and Tomás Rau. 2012. "Tournament Incentives for Teachers: Evidence from a Scaled-up Intervention in Chile." *Economic Development and Cultural Change* 61(1): 219–246.

Cross, Michael, Ratshi Mungadi, and Sepi Rouhani. 2002. "From Policy to Practice: Curriculum Reform in South African Education." *Comparative Education* 38(2): 171–187.

Crouch, Luis. 2005. "Political Economy, Incentives, and Teachers' Unions." In *Incentives to Improve Teaching*: Lessons from Latin America, edited by Emiliana Vegas, 389. Washington, DC: World Bank.

Cuenca, Ricardo. 2017. *Moving toward Professional Development. The Teacher Reform in Peru (2012–2016).* Instituto de Estudios Peruanos.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2): 31–47.

Dahlum, Sirianne, and Carl H. Knutsen. 2017. "Do Democracies Provide Better Education? Revisiting the Democracy-Human Capital Link." *World Development* 94: 186–199.

De Ree, Joppe, Karthik Muralidharan, Menno Pradhan, and Halsey Rogers. 2015. *Double for Nothing? Experimental Evidence on the Impact of an Unconditional Teacher Salary Increase on Student Performance in Indonesia.* Working paper no. 21806. National Bureau of Economic Research.

Delannoy, Françoise. 2000. *Educational Reforms in Chile, 1980–1998: A Lesson in Pragmatism.* Washington, DC: World Bank.

Department of Basic Education. 2018. *National Curriculum Statements (NCS) Grades R–12*. Republic of South Africa. https://www.education.gov.za/Curriculum/NationalCurriculumStatementsGradesR–12.aspx.

Di John, Jonathan, and James Putzel. 2009. *Political Settlements: Issues Paper*. Discussion paper. University of Birmingham.

Draibe, Sônia. 2004. "Federal Leverage in a Decentralized System: Education Reform in Brazil." In *Crucial Needs, Weak Incentives: Social Sector Reform, Democratization, and Globalization in Latin America*, edited by Robert R. Kaufman and Joan M. Nelson, 375–406. Baltimore: Johns Hopkins University Press.

Drèze, Jean, and Amartya Sen. 1995. *India: Economic Development and Social Opportunity.* Delhi; New York: Oxford University Press.

Ferraz, Claudio, Frederico Finan, and Diana B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96(9–10): 712–726.

Fierro, Margarita Z., and José M. Velasco. 2002. "Ten Years after the National Accord for the Modernization of Basic Education in Mexico: Challenges, Tensions and Perspectives." *Revista Electrónica de Investigación Educativa* 4(2): 2.

Filmer, Deon, and Lant Pritchett. 1999. "The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries." *Population and Development Review* 25(1): 85–120.

Firpo, Sergio, Renan Pieri, and André P. Souza. 2017. "Electoral Impacts of Uncovering Public School Quality: Evidence from Brazilian Municipalities." *EconomiA* 18(1): 1–17.

Gershberg, Alec I. 1999. "Education 'Decentralization' Processes in Mexico and Nicaragua: Legislative Versus Ministry-Led Reform Strategies." *Comparative Education* 35(1): 63–80.

Gómez-Zaldívar, Manuel. 2014. "Has the National Agreement for the Modernization of Basic Education Contributed to Improving Levels of Basic Education and Reducing Disparities between the States?" *Latin American Journal of Economics* 51(1): 85–107.

Govender, Logan. 2015. "Teacher Unions' Participation in Policy Making: A South African Case Study." *Compare: A Journal of Comparative and International Education* 45(2): 184–205.

Government of the United Republic of Tanzania. 2013. *Big Results Now! Roadmap*. Tanzania Development Vision 2025.

Grindle, Merilee S. 2002. *First in the Queue? Mainstreaming the Poor in Service Delivery*. Paper presented at World Development Report 2003/04 workshop held at Eynsham Hall, Oxford.

Grindle, Merilee S. 2004. *Despite the Odds: The Contentious Politics of Education Reform.* Princeton: Princeton University Press.

Hanushek, Eric A., and Ludger Woessmann. 2007. *The Role of Education Quality for Economic Growth.* Policy research working paper series no. 4122. Washington, DC: World Bank. http://elibrary.worldbank.org/doi/abs/10.1596/1813–9450–4122.

Harding, Robin, and David Stasavage. 2013. "What Democracy Does (and Doesn't Do) for Basic Services: School Fees, School Inputs, and African Elections." *Journal of Politics* 76(1): 229–245.

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik. 2005. "Growth Accelerations." *Journal of Economic Growth* 10(4): 303–329.

Hecock, R. Douglas. 2014. "Democratization, Education Reform, and the Mexican Teachers' Union." *Latin American Research Review*: 62–82.

Hickey, Samuel, Kunal Sen, and Badru Bukenya. 2015. "Exploring the Politics of Inclusive Development: Towards a New Conceptual Approach." In *The Politics of Inclusive Development: Interrogating the Evidence*, edited by Sam Hickey, Kunal Sen, and Badru Bukenya, 3–32. Oxford University Press: Oxford.

Hickey, Samuel, and Naomi Hossain, 2019. *The Politics of Education in Developing Countries: From Schooling to Learning*. Oxford: Oxford University Press.

Hossain, Naomi, and Mick Moore. 2002. *Arguing for the Poor: Elites and Poverty in Developing Countries.* Brighton: Institute of Development Studies.

Hossain, Naomi, Mirza Hassan, Muhammad A. Rahman, Khondokar S. Ali, and M. Sajidul Islam. 2017. *The Problem with Teachers: The Political Settlement and Education Quality Reforms in Bangladesh.* ESID working paper no. 86. University of Manchester.

International Commission on Financing Global Education Opportunity (ICFGEO). 2017. *The Learning Generation: Investing in Education for a Changing World.* New York: ICFGEO.

James, Selma. 2014. "What We Can Learn from Tanzania's Hidden Socialist History." *The Guardian*. https://www.theguardian.com/commentisfree/2014/dec/11/tanzania-hidden-socialist-history-president-julius-nyerere.

Jansen, Jonathan D. 1998. "Curriculum Reform in South Africa: A Critical Analysis of Outcomes-Based Education." *Cambridge Journal of Education* 28(3): 321–331.

Khan, Mushtaq. 2010. "*Political Settlements and the Governance of Growth-Enhancing Institutions. Draft Paper.*" Research paper series on "Growth-Enhancing Governance." SOAS University of London.

Kingdon, Geeta G., Angela Little, Monazza Aslam, Shenila Rawal, Terry Moe, Harry Patrinos, Tara Beteille, Rukmini Banerji, Brent Parton, and Shailendra K. Sharma. 2014. *A Rigorous Review of the Political Economy of Education Systems in Developing Countries.* London: EPPI-Centre, University of London.

Kosack, Stephen. 2012. *The Education of Nations: How the Political Organization of the Poor, Not Democracy, Led Governments to Invest in Mass Education.* Oxford: Oxford University Press.

Kraak, Andre, and Michael F.D. Young. 2001. *Education in Retrospect: Policy and Implementation Since 1990.* Pretoria: HSRC Press.

Laterite. 2017. *Assessment of Repetition and Dropout in Basic Education in Rwanda.* Kigali: UNICEF.

Levy, Brian. 2014. *Working with the Grain: Integrating Governance and Growth in Development Strategies.* Oxford: Oxford University Press.

Levy, Brian, Robert Cameron, Ursula Hoadley, and Vinothan Naidoo. 2018. *The Politics and Governance of Basic Education: A Tale of Two South African Provinces.* Oxford: Oxford University Press.

Machin, Stephen, and Sandra McNally. 2004. *Large Benefits, Low Cost: Is the Government's National Literacy Strategy Effective?* London: CentrePiece, London School of Economics.

———. 2008. "The Literacy Hour." *Journal of Public Economics* 92(5): 1441–1462.

Mani, Anandi, and Sharun Mukand. 2007. "Democracy, Visibility and Public Good Provision." *Journal of Development Economics* 83(2): 506–529.

Martin, Eric. 2017. "Mexican President's Support Plumbs New Low As Gasoline Soars." *Bloomberg*. https://www.bloomberg.com/news/articles/2017–01–18/mexico-s-pena-nieto-approval-falls-to–12-after-gasoline-soars.

McEwan, Patrick, and Lucrecia Santibáñez. 2005. *Teacher and Principal Incentives in Mexico*. Brief no. 81470. Washington, DC: World Bank.

McLoughlin, Claire. 2011. "Factors Affecting State-Non-Governmental Organisation Relations in Service Provision: Key Themes from the Literature." *Public Administration and Development* 31(4): 240–251.

McNally, Sandra. 2015. "Numeracy and Mathematics." In *Making a Difference in Education: What the Evidence Says,* edited by Robert Cassen, Sandra McNally, and A. Vignoles, 123–134. London: Routledge.

Metzler, Johannes, and Ludger Woessmann. 2012. "The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from within-Teacher within-Student Variation." *Journal of Development Economics* 99(2): 486–496.

Mizala, Alejandra, and Ben R. Schneider. 2014. "Negotiating Education Reform: Teacher Evaluations and Incentives in Chile (1990–2010)." *Governance* 27(1): 87–109.

Mollel, Henry A., and Albertjan Tollenaar. 2013. "Decentralization in Tanzania: Design and Application in Planning Decisions." *International Journal of Public Administration* 36(5): 344–353.

MPME. 2013. *National Student Assessment 2013 for Grades 3 and 5.* Dhaka: Bangladesh Ministry of Primary and Mass Education.

Msibi, Thabo, and Sihle Mchunu. 2013. "The Knot of Curriculum and Teacher Professionalism in Post-Apartheid South Africa." *Education As Change* 17(1): 19–35.

Mullis, Ina V.S., Michael O. Martin, Pierre Foy, and Martin Hooper. 2016. *TIMSS 2015 International Results in Mathematics.* TIMSS & PIRLS International Study Center at Boston College.

Murillo, Maria V. 1999. "Recovering Political Dynamics: Teachers' Unions and the Decentralization of Education in Argentina and Mexico." *Journal of Interamerican Studies and World Affairs* 41(1): 31–57.

Nath, Samir R., and Ahmed M.R. Chowdhury. 2009. *State of Primary Education in Bangladesh: Progress Made, Challenges Remain.* Dhaka: Campaign for Popular Education.

NeuroMat. 2017. "An Overview of Mathematical Literacy in Brazil: An Interview with PISA Analyst Esther Carvalhaes." http://neuromat.numec.prp.usp.br/content/overview-mathematical-literacy-brazil-interview-pisa-analyst-esther-carvalhaes.

NKRA. 2014. *Education NKRA Lab Report: Executive Summary.* Development Partners Group Tanzania. http://www.tzdpg.or.tz/fileadmin/documents/dpg_internal/dpg_working_groups_clusters/cluster_1/psdtrade/Documents/Policy_docs/2._Education_NKRA_-_Executive_Summary.pdf.

———. 2016. *National Key Result Area 2013/14–2015/16: Education.* Tanzania: Minister of Education.

OECD. 2008. *Reviews of National Policies for Education: South Africa.* Paris: OECD Publishing.

———. 2010. *Education at a Glance 2010: OECD Indicators.* Paris: OECD Publishing.

———. 2016. *PISA 2015 Results (Volume I).* Paris: OECD Publishing.

Ofsted. 1996. *The Teaching of Reading in 45 Inner London Primary Schools.* London: Ofsted.

Ornelas, Carlos. 2000. "The Politics of the Educational Decentralization in Mexico." *Journal of Educational Administration* 38(5): 426–442.

———. 2004. "The Politics of Privatisation, Decentralisation and Education Reform in Mexico." *International Review of Education* 50(3–4): 397–418.

———. 2006. "The Politics of Privatisation, Decentralisation and Education Reform in Mexico." In *Decentralisation and Privatisation in Education,* edited by Joseph Zajda, 207–228. Dordrecht: Springer.

Ornelas, Nancy Otero, Gabriela Calderon, and Paulo Blikstein. 2014. *Makers in Residence Mexico: Creating the Conditions for Invention.* Short paper at FabLearn Europe: Digital Fabrication in Education Conference.

Parandekar, Suhas, Érica Amorim, and Andrea Welsh. 2008. *Prova Brasil: Building a Framework to Promote Measurable Progress in Learning Outcomes.* Brief no. 121. Washington, DC: World Bank.

Patrinos, Harry A. 2015. *¿El Suministro De Información Sobre El Desempeño De Las Escuelas Puede Conducir a Un Mejoramiento?* Banco Mundial Blogs. http://blogs.worldbank.org/voices/es/el-suministro-de-informacion-sobre-el-desempeno-de-las-escuelas-puede-conducir-un-mejoramiento.

Patton, Michael Q. 2002. *Qualitative Research and Evaluation Methods.* London: SAGE.

Pont, Beatriz, Guillermo M., Etienne Albiser, Juliana Zapata, and Sylvain Fraccola. 2013. *Education Policy Outlook: Chile.* Paris: OECD Publishing.

Prensa Latina. 2017. "Mexican Teachers Return to Protest the Educational Reform." Agencia Informativa Latino Americana Prensa Latina. https://www.plenglish.com/index.php?o=rn&id=8006&SEO=mexican-teachers-return-to-protest-the-educational-reform.

Pritchett, Lant. 2013. *The Rebirth of Education: Schooling Ain't Learning.* Washington, DC: Center for Global Development.

Puttick, Ruth, Peter Baeck, and Philip Colligan. 2014. *The Teams and Funds Making Innovation Happen in Governments around the World.* Nesta. http://theiteams.org/system/files_force/i-teams_June%202014.pdf#page=74.

Ramatlapana, Kim, and Judah P. Makonye. 2012. "From Too Much Freedom to Too Much Restriction: The Case of Teacher Autonomy from National Curriculum Statement (NCS) to Curriculum and Assessment Statement (CAPS)." *Africa Education Review* 9(sup1): S7–S25.

Reddy, Vijay, Tia L. Zuze, Mariette Visser, Lolita Winnaar, Andrea Juan, Cas Prinsloo, Fabian Arends, and Shawn Rogers. 2015. *Beyond Benchmarks: What Twenty Years of TIMSS Data Tell Us About South African Education.* Pretoria: HSRC Press.

Richards, John, and Aidan R Vining. 2016. "Achieving Better Primary School Outcomes in a Context of Weak Education Governance: A Case Study of Bangladesh." *Bangladesh Education Journal*, 15(2).

Rieffel, Lex. 2007. *Indonesia: Ten Years after the Crisis*. Brookings Op-Ed. https://www.brookings.edu/opinions/indonesia-ten-years-after-the-crisis/.

RISE. n.d. *Big Results Now! Research on Education System Reform in Tanzania: Technical Proposal.* Research on Improving Systems of Education.

Rivero, José. 2010. *Teacher Performance and Its Impact on Quality Education in Peru: Policy Paper*. FOCAL, Canadian International Development Agency.

Rosser, Andrew, and Mohamad Fahmi. 2016. *The Political Economy of Teacher Management in Decentralized Indonesia*. Policy research working paper. Washington, DC: World Bank Group.

Sabarwal, Shwetlena, Arun Joshi, and William Blackmon. 2016. *A Review of Multi Donor Results Based Financing Mechanism Used for Big Results Now in Education Program.* Washington, DC: World Bank.

SACMEQ. 2018. *The Southern and Eastern Africa Consortium for Monitoring Educational Quality*. University of Botswana. http://www.sacmeq.org/?q=sacmeq-projects/sacmeq-i/readingmathscores.

SACMEQ III. 2017. *The Southern and Eastern Africa Consortium for Monitoring Educational Quality*. SACMEQ. http://www.sacmeq.org.

Sandefur, Justin, and Amanda Glassman. 2015. "The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics." *Journal of Development Studies* 51(2): 116–132.

Santibañez, Lucrecia, and Brenda J. Rabling. 2007. "Conflict and Power: The Teacher's Union and Education Quality in Mexico." *Well-Being and Social Policy* 3(2): 21–40.

Sriprakash, Arathi, Leon Tikly, and Sharon Walker. 2019. "The Erasures of Racism in Education and International Development: Re-Reading the 'Global Learning Crisis.'" *Compare: A Journal of Comparative and International Education*.

Stannard, John, and Laura Huxford. 2007. *The Literacy Game: The Story of the National Literacy Strategy.* London: Routledge.

Stasavage, David. 2005. "The Role of Democracy in Uganda's Move to Universal Primary Education." *Journal of Modern African Studies* 43(1): 53–73.

Sumra, Suleman, and Joviter Katabaro. 2014. *Declining Quality of Education: Suggestions for Arresting and Reversing the Trend.* Economic and Social Research Foundation.

Tanner, Emily, Ashley Brown, Naomi Day, Mehul Kotecha, Natalie Law, Gareth Morrell, Ola Turczuk, Victoria Brown, Aleks Collingwood, and Haroon Chowdry. 2010. *Evaluation of Every Child a Reader (ECaR).* Research report DFE-RR114.

TIMSS. 2015. *Trends in International Mathematics and Science Study 2015.* http://timssandpirls.bc.edu/.

Tymms, Peter. 2004. "Are Standards Rising in English Primary Schools?" *British Educational Research Journal* 30(4): 477–494.

Ulrich, Dave, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics* 1(1): 112–135.

UNESCO. 2012. *UNESCO Institute for Statistics*. http://data.uis.unesco.org/.

———. 2013. *EFA Global Monitoring Report 2013/14: Teaching and Learning: Achieving Quality for All.* Paris: UNESCO.

———. 2014. *Teaching and Learning: Achieving Quality for All, EFA Global Monitoring Report, 11.2013/14.* Paris: UNESCO.

———. 2017. *UNESCO Institute for Statistics*. http://data.uis.unesco.org/.

Uwezo. 2012. *Are Our Children Learning? Annual Learning Assessment Report 2012.* Dar Es Salaam: Uwezo Tanzania. http://www.uwezo.net/wp-content/uploads/2012/08/TZ_Uwezo2012ALAReport.pdf.

———. 2014. *Are Our Children Learning? Literacy and Numeracy across East Africa 2013.* Nairobi: Uwezo East Africa at Twaweza.

Williams, Timothy P. 2016. *Oriented towards Action: The Political Economy of Primary Education in Rwanda.* Manchester, UK: University of Manchester.

———. 2017. "The Political Economy of Primary Education: Lessons from Rwanda." *World Development* 96: 550–561.

———. 2019. "The Downsides of Dominance: Education Quality Reforms and Rwanda's Political Settlement." In *The Politics of Education in Developing Countries*, edited by Samuel Hickey and Naomi Hossain, 86–104. Oxford: Oxford University Press.

World Bank. 2013a. *Project Appraisal Document for a Third Primary Education Development Program.* Washington, DC: World Bank.

———. 2013b. *Seeding Fertile Ground: Education that Works for Bangladesh.* Washington, DC: World Bank Group.

———. 2014. "*Project Appraisal Document for the Big Reults Now in Education (BRNEd) Program for the Republic of Tanzania.*" Washington, DC: World Bank.

———. 2016. *Tanzania Big Results Now in Education Program: Mid-Term Review.* Washington, DC: World Bank.

———. 2017. *World Development Indicators*. https://datacatalog.worldbank.org/dataset/world-development-indicators.

———. 2018. *The World Development Report 2018: Learning to Realize Education's Promise.* Washington DC: World Bank.

Yin, Robert K. 2009. *Case Study Research: Design and Methods.* London: SAGE.

Zorrilla Fierro, Margarita. 2002. "Diez Años Después Del Acuerdo Nacional Para La Modernización De La Educación Básica En México: Retos, Tensiones Y Perspectivas." *Revista electrónica de investigación educativa* 4(2): 1–19.

Annex A: Approach to identify episodes of sustained change in learning outcomes

This section outlines the approach taken to identify episodes of change—growth or decline—in the PISA reading scores.

1. Step 1: Drop countries if they have
    a. participated in only one year (e.g., Panama) or
    b. participated in only two consecutive test years (e.g., Vietnam, Azerbaijan).

2. Step 2: Create overlapping intervals of six-year periods [2000–2006, 2003–2009, 2006–2012, 2009–2015].
    a. For countries with gaps in participation, some intervals are greater than six years (e.g., for Peru, which did not participate in 2003 and 2006, the first interval equals 2000–2009).
    b. For other countries with participation gaps, consecutive intervals may not overlap (e.g., for Bulgaria, which did not participate in 2003, the first interval equals 2000–2006 and second interval equals 2006–2012).

3. Step 3: Identify the minimum change in scores needed to count as a significant change from one year to another for each country.
    a. For each country and test year, calculate how many points scores need to change to detect a statistically significant change at the 1% level. This value (Min_{ct}) for country *c* in year *t* is calculated as follows,

        $Min_{it} = |2.58*SE_{ct}|$, where SE_{ct} denotes the standard error of the mean score for country *c* in year *t*.
    b. For each country, take the average of the minimum scores identified in (a) across all test years. This average will serve as a benchmark to determine whether scores between two test years changed significantly or not.

4. Step 4: For each interval created in Step 2, calculate the change in scores between the beginning and end of the interval, and categorize as follows:
    a. Let MIN_i denote the minimum score change identified in Step 3 (b) for country *c*.
    b. If absolute change < MIN_c, code interval trend as "no significant change."
    c. If change is positive, monotonic, and >= MIN_c, code interval trend as "increasing."
    d. If change is negative, monotonic, and absolute change >= MIN_c, code interval trend as "decreasing."
    e. For nonmonotonic (significant) **increase** over the interval, adopt the following rules:
        i. Let ***delta1*** denote change between first two consecutive years in each interval, and ***delta2*** denote change between second two consecutive years in the interval.
        ii. If delta1 is negative and delta2 is positive, code interval as increasing if delta1 is insignificant (i.e., |delta1| < MIN_c); otherwise code interval as "undefined."
        iii. If delta1 is positive and delta2 is negative, code interval as increasing if delta2 is insignificant; otherwise code interval as "undefined."
    f. For nonmonotonic (significant) **decrease** over the interval, adopt the following rules:

i. If delta1 is negative and delta2 is positive, code interval as decreasing if delta2 is insignificant; otherwise code as "undefined."
ii. If delta1 is positive and delta2 is negative, code interval as decreasing if delta1 is insignificant; otherwise code as "undefined."

5. Step 5: For each interval, check if the interval trend can be extended to years following or preceding the interval. For example, if interval from 2006 to 2012 for Thailand is identified as increasing, can the end point of this increasing interval be plausibly extended to 2015? Can the starting point for this interval be extended back to 2003 (or to 2000)? Adopt the following rules to determine whether and how far intervals created in Step 2 can be extended.
    a. Let $score_j$ denote score in the first year of an interval; let $score_{j-1}$ denote score in the first test year preceding the interval and so on.
    b. Let $score_k$ denote score in the last year of an interval; let $score_{k+1}$ denote score in the first test year following the interval and so on.
    c. If interval *i* for country *c* is coded as increasing in Step 4 and
        i. $[score_{k+1} - score_k] > 0$, annex $year_{k+1}$ to the interval;
        ii. $[score_{k+1} - score_k] <= 0$ and $|score_{k+1} - score_k| < MIN_c$ and $[score_{k+1} - score_j] >= MIN_c$, annex $year_{k+1}$ to the interval (i.e., if the score in the year following the interval is lower than score in the end of the interval, but the change is insignificant, annex the year to the interval as long as the score in $year_{k+1}$ is significantly higher than the score in the beginning of the interval);
        iii. $[score_{k+1} - score_k] > 0$, repeat steps (i), (ii), and (iii) for the second and third years following the end of the interval;
        iv. $[score_j - score_{j-1}] > 0$, annex $year_{j-1}$ to the interval;
        v. $[score_j - score_{j-1}] <= 0$ and $| score_j - score_{j-1}| < MIN_c$ and $[score_k - score_{j-1}] >= MIN_c$, annex $year_{j-1}$ to the interval (if interval was extended to include year(s) following the end of the original interval, $score_k$ denotes the score in the end of the new extended interval); and
        vi. $[score_j - score_{j-1}] > 0$, repeat steps (iv), (v), and (vi) for the second and third years following the end of the interval.

    d. If interval *i* for country *c* is coded as decreasing in Step 4, and
        i. $[score_{k+1} - score_k] < 0$, annex $year_{k+1}$ to the interval;
        ii. $[score_{k+1} - score_k] >= 0$ and $|score_{k+1} - score_k| < MIN_c$ and $|score_{k+1} - score_j| >= MIN_c$; annex $year_{k+1}$ to the interval (i.e., score in year following the interval is higher than score in the end of the interval, but the change is insignificant, annex the year to the interval as long as the score in $year_{k+1}$ is significantly lower than the score in the beginning of the interval);
        iii. $[score_{k+1} - score_k] < 0$, repeat steps (i), (ii), and (iii) for the second and third years following the end of the interval;
        iv. $[score_j - score_{j-1}] < 0$, annex $year_{j-1}$ to the interval;
        v. $[score_j - score_{j-1}] >= 0$ and $| score_j - score_{j-1}| < MIN_c$ and $|score_k - score_{j-1}| >= MIN_c$, annex $year_{j-1}$ to the interval (if interval was extended to include year(s) following the end of the original interval, $score_k$ denotes the score in the end of the new extended interval); and

vi. $[score_j - score_{j-1}] < 0$, repeat steps (iv), (v), and (vi) for the second and third years following the end of the interval.

Table A1: Average number of activities introduced under reform category per reform period

Panel A: High-income countries	Episodes of decline	Episodes of improvement
Ensuring equity in access and retention	4.9	3.9
Curriculum reforms	3.2	3
Improving teacher quality	2.5	2.7
Structural reforms	2.5	2.3
ICT use in education	1.9	2.6
Learning assessment systems	2	2
School-based management	2	1.7
Skills development	2.6	2
Building institutional capacity	2.8	2.8
Provision of textbooks, learning materials, or free services	1.8	1.2
School/classroom construction for expanding access	2	1.8
Hiring more teachers	1	1
Teacher pay-and-performance evaluations	1.4	1.7
Other approaches to learning (besides ICT use)	1	1.4
Panel B: Low- and middle-income countries		
Improving teacher quality	2.9	5
Building institutional capacity	3.5	5.1
Ensuring equity in access and retention	3.1	5.7
Curriculum reforms	1.5	3.1
School-based management	0.8	2.6
Structural reforms	2.4	3
Provision of textbooks, learning materials or free services	1.4	2.7
Learning assessment systems	1	2.4
Skills development	2	2.1
School/classroom construction for expanding access	1.1	2.1
ICT use in education	0.6	2.3
Teacher pay-and-performance evaluations	0.2	1.4
Hiring more teachers	0	2
Other approaches to learning (besides ICT use)	0	1.3

Notes: Average number of activities are based on the episodes during which at least one activity under the broader reform category was introduced. ICT stands for information and communications technology.

Annex B: Country case studies

Country case 1: Improving funding and monitoring in Brazil

Introduction

1. This case provides a review of learning trends, reforms in education, and the broader political context from 1995 to 2015. Brazil has experienced an improvement in learning outcomes, and education is becoming more equitable. The last 20 years saw major expansion and improvements in quality, especially in primary levels. Brazil has transformed from one of the lowest performers in the PISA 2000 to having sustained improvement. There are three notable reforms that have been instrumental in bringing qualitative improvements to the education system.

   The first reform was municipal decentralization in the mid-to-late 1990s.[80] It served as a vehicle to extend federal resources based on a redistributive policy. The initial focus was on improving primary education and entrusting education management to state and municipal governments.[81] The Primary Education Fund, also known as the Fund for Maintenance and Development of Basic Education and Valuation of Education Professionals (FUNDEF, later known as FUNDEB), helped expand student coverage and local accountability.

   The second reform was in 1995, when the government developed assessments and formed a results framework system with the System for Evaluation of Basic Education (SAEB). In 2005, the government extended SAEB into a census-based learning assessment, renamed Prova Brasil. The third reform was when the government introduced the Basic Education Development Index (IDEB), a monitoring tool that built upon reforms grounded in the 1990s. IDEB is a major success in itself, and it paved the way to consolidating several reform initiatives.

   Decisive actors of reform came from within Brazil's political system.[82] This case investigates how, through good leadership and local coalition building, Brazil has improved education though successful reforms. This case provides a short overview of learning trends, the broader political context during the reform period, a description of the reforms, motivations for reform, the reform process, and lessons learned.

Data/evidence: Trends in learning outcomes and definition of reform episode/period

2. Both regional and national assessments have shown a modest, but overall positive, trend in learning outcomes. Brazil's education system has gradually become more inclusive 2000

[80] Lei de Diretrizes e Bases da Educação Nacional, the National Education Foundations, and the Guidelines Act of 1996 also granted states and municipalities greater autonomy in preparing curricula; see Andrews and Vries (2012).
[81] Starting with the 1988 constitution, Brazil progressively decentralized funding for basic education; see Parandekar, Amorim, and Welsh (2008).
[82] Draibe (2004).

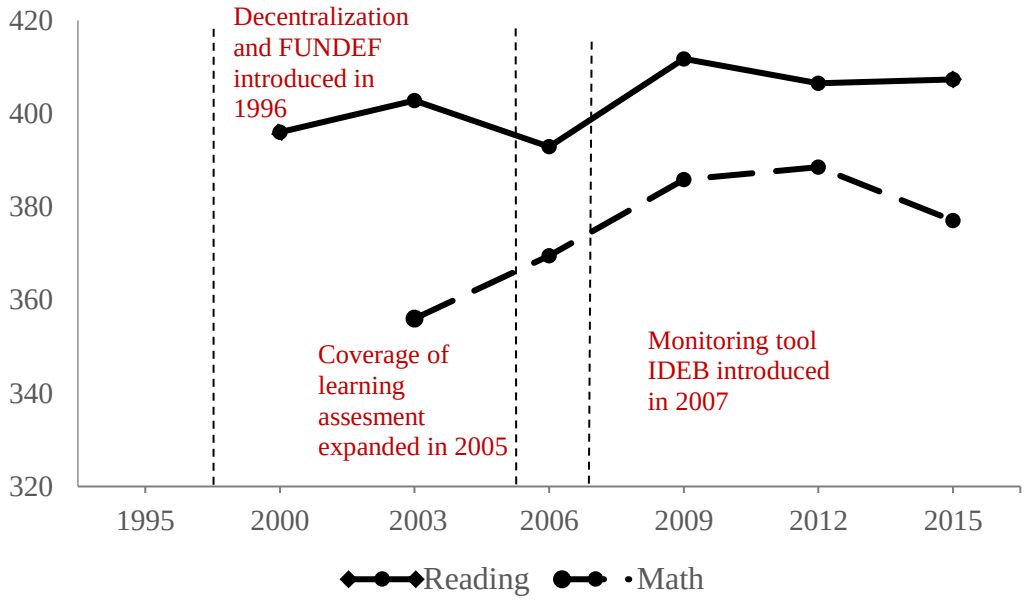
onward (net enrollment since 2000 is, on average, 95% in primary, and as of 2017 is 78% in secondary).[83] There have been improvements in learning outcomes, not only in average performance levels but also with disadvantaged students who made gains in learning. The largest increase in primary school coverage between 1997–2000 was in municipal schools in the northeast and north, the poorest region.[84] By 2009, the nine-year compulsory schooling cycle started, and it became one of the longest in the region,[85] which may have also helped increase enrollment.

3. There is an overall positive trend in learning. Results in national, and especially international, learning assessments reinforce this finding. PISA math scores saw a sharp increase mid-2000–2012 of 30 points, and both disadvantaged and advantaged students increased their performance on the PISA test.[86] Based on analyses of reading and math scores for both groups, Carnoy et al. argue, however, that the most plausible reason for overall gains in early PISA results is largely due to the significant increase in years of schooling, not the greater efficiency of math teaching in each grade.[87] Nevertheless, "Brazilian students' gains in mathematics are as high as or higher than any comparison country students' gains except Peru's."[88] Scores from national assessments conducted at the primary level also show better math performance starting in the early 2000s.

4. The 2006–2015 period marked increasing trends in PISA reading scores (Figure A1). This was a big improvement compared to the mid-2000s; from 403 in 2003, down to 393 in 2006, and up to 412 in 2009. In science, from 2006 to 2012, gains across performers were similar: both the 10th and 90th percentile scores increased by 21 points. The top performers made gains from 2012 to 2015 (17 points), but the bottom 10% of performers' scores declined by 12 points.

[83] UNESCO (2017).
[84] Draibe (2004).
[85] In 2006, the country adopted legislation extending the length of compulsory schooling by one year; see
[86] Results are based on Klein's method of weighting scores (2000–2012) by grade for each of the two family academic resources (FAR) groups (0–10 books in the home and > 100 BH); see Carnoy et al. (2015).
[87] Carnoy et al. (2015).
[88] Latin American countries, Spain, and Portugal; see Carnoy et al. (2015).

Figure A1: Mean scores in PISA reading and math over time



Source: OECD (2016).
Notes: FUNDEF stands for Fund for Maintenance and Development of Basic Education and Valuation of Education Professionals and IDEB for Basic Education Development Index.

5. Equity in education has begun to emerge (2000–2015), as students with fewer resources are gaining on advantaged students. Education access rose and national learning trends became positive.[89] The math assessments, SAEB, and Prova Brasil track learning levels of students in 4th, 8th, and 11th grades. Students with fewer resources in the home are definitely not falling farther behind their more advantaged counterparts. In fact, they may be superseding them. Results are based on measures of FAR,[90] an assessment of whether a country's educational system is improving, going beyond only tracking average national scores. The significant gains in math, and the smaller gains in Portuguese, were concentrated among students in low and middle FAR groups.[91]

6. From a regional perspective, improved performance in learning varied, even for disadvantaged groups. By 2012, results showed that disadvantaged Brazilian students were not scoring as high in math and reading as their counterparts in Latin American countries, except in the case of Argentina.[92] Yet Brazilian students at all levels of FAR performed as well or better than students in Argentina, Colombia, and Peru but not as well as students in Chile, Mexico, Uruguay, Portugal, or Spain.[93]

[89] Based on measures in mother's education (ME) data; see Carnoy et al. (2015).
[90] Proxy measures of students' FAR, such as ME, parents' education, articles in the home, or books in the home, are compared to students' academic achievement.
[91] Such results were found using two different measures of FAR—the number of books at home and ME; see Carnoy et al. (2015).
[92] Carnoy et al. (2015).
[93] Carnoy et al. (2015).

Broader context: Reforms spanning two decades (1995–2015)

7. The most successful reforms can be seen in two waves over time. Known as the "silent revolution" in lower education echelons, the first wave ushered in pioneering. Strong, central leadership roles spearheaded the reform process. The Cardoso government (1996–2003) introduced radical education reforms, and the consecutive progressive administrations of Lula da Silva (2002) and Dilma Rousseff (2011) built upon reforms that aimed to expand coverage and improve quality. The second reform wave is the process of reform consolidation and strengthens access toward a quality education system, especially 2005 onward. Throughout this process of change, coalitions and partnerships helped fuel reform incentives.

8. Before the reform period, the Brazilian education system was in a state of inertia, and most reforms were phased in gradually. Basic education reform began in 1971 under military rule, yet the Ministry of Education was a "center of patronage and clientelism," one that failed to introduce innovation in the early part of that decade.[94] Both then and during the democratization process of 1986, a top-down power system existed. Influence for reform emanated from the country's political elite, where stakeholder feedback was limited.[95] State and municipal administrations depended upon the federal government, which had regulatory and economic power. Before 2000, mobilization in the education sector did not exist, but social mobilization pushing for education reforms gained real momentum in the 2000s.

9. In 1997, reforms gained real momentum when the government introduced FUNDEF. It came in a context of hyperinflation across the region and a period of stringent fiscal adjustment, yet the fund enabled municipal spending incentives and allowed for greater impact at the school level in smaller disadvantaged municipalities. It also led to an increase in education spending, from 2% of GDP in 1995 to 4% in 2008.[96] It is important to mention that in 1971, the regime allocated more revenue toward education through tax revenues, and the term "educational wages"[97] had major significance on later reforms. This "wage" is a compulsory 2.5% firm payroll tax that all firms must pay, allocating it to the Education Ministry spending (especially earmarked for the Maintenance of the Development of Teaching (PMDE), discussed next).[98]

10. By 2010, Brazil had established political and economic stability. The country is still largely driven by commodities and relies on unskilled labor in many sectors. But in other sectors of the economy, an educated, skilled workforce is also required (i.e., in highly innovative and internationally competitive companies in aeronautical, petrochemicals, natural gas, mining, steel, paper and pulp, ethanol, and meat). The workforce averages nine years of education, with on-the-job skill provision. In 1990, the average schooling level of the labor force was 3.8 years, less than half that of Argentina, Chile, and the OECD.

[94] Draibe (2004).
[95] Draibe (2004).
[96] Bruns, Evans, and Luque (2011).
[97] A constitutional norm known as the Calmon Amendment, 1983.
[98] Under the Medici and later Figueiredo Administration; see Draibe (2004).

11. In this context, reform accomplishments come with its challenges. In terms of education quantity, the challenge has been to improve quality, provide more years of schooling, and improve high school completion rates (also a quality indicator). Quality wise, Brazil's case shows ambition on how a low-performing country has steadily increased learning outcomes using new performance benchmarks. It has also adopted best practices in teacher quality and strong school management, improved instructional systems, and provided better teacher preparation, qualifications, salaries, and training. Through this process, Brazil has become an inspiring model of state-led educational development for other countries.

Education reform program

12. The education resource distribution (FUNDEF) and assessments (SAEB/ProvaBrazil, IDEB) helped found Brazil's education framework. A new democratic administration began political bargaining and coalition building around reforms.[99] The first wave of changes coincided with the newly elected President Fernando Cardoso, the first lady, and the Minister of Education, Renato de Souza, who led the reform process. Their reform efforts have helped bring about considerable accountability. The initial movement of the 1990s focused on reducing allocative inequalities, combating student repetition, and improving enrollment and average per-student spending. And in the 2000s, a concerted focus on quality learning and student retention emerged.

13. Many efforts to improve quality have since taken shape due being established of these reforms. Education resource decentralization, better known as "municipalization" (still ongoing) has given agency to subnational governments and school directors.[100] The fund had catalyzed Cardoso's administration to reorganize institutions toward "municipalizing revenues."[101] In other words, the administration aimed to decentralize federal resources and programs that support state and municipal schools. Cardoso had brought on board key leadership from the PMDE team (founded in 1995). Minister Souza gained presidential backing for the early School Lunch Program and expanded provisions. He also undertook various institutional reforms, especially the decentralization of the budget and the reform of the teachers' code. Hence, he became known as a "skillful negotiator."[102]

14. Incentives improved as enrollment rates rose with FUNDEB resources (2005–2010). The decentralization process complemented this expansion of provisions. By 2004, rural investments included secondary school instructional and learning material and transport and meals (the daily student value was adjusted by 38% for the first time since 1995), reaching

[99] In the Brazilian educational context there are two prominent coalitions, the "Campanha Nacional pelo Direito à Educação" and the "Todos pela Educação."
[100] Draibe (2004).
[101] Nonetheless, adverse effects were detected in cases where corruption involved educational block grants transferred from the central government to municipalities. Learning had a significant negative link to corruption and the school performance of primary school students; see Ferraz, Finan, and Moreira (2012).
[102] Draibe (2004).

around 3.2 million rural students.[103] The guaranteed minimum per-pupil allocation meant "that funding would follow the student, which created a significant incentive for school systems—especially underfunded municipal systems—to expand their enrollments."[104]

15. The fund powered equity in accessing basic education. Draibe explains that the government channeled resources from wealthier state schools to poorer municipal schools, and this had a redistributive effect.[105] By 2010, for instance, the FUNDEB allocated 25% of state revenue toward education spending.[106] The addition of federal funds to poor states has provided schools resources and made them more comparable to those available to wealthier states' schools. States that fell below minimum expenditure standards could be federally compensated through FUNDEB. Teacher salaries rose, on average, 13%, and in the poor northeast it rose to 60%.[107] The extra federal funding helped give poor states resources and ensured that all elementary schools had a minimum per-pupil allocation.

16. In 1995 the creation of a national evaluation system (SAEB) helped track learning and raise incentives. [Prova Brazil and IDEB] "are the measurement anchor for a new wave of policies in Brazil aimed at creating stronger incentives for teachers and schools."[108] Data platforms can also ensure that the cost of acquiring information and monitoring is low. The SAEB[109] is a sample-based learning assessment that later became census based. Developed in 2015, it was later renamed Prova Brazil. The assessment design allowed it to be comparable across years.

    In 2007, IDEB was also set up to solidify accountability.[110] Customized IDEB targets were set up for each state, municipality, and district school. This internationally benchmarked system gauges not only individual schools but also municipal and state systems. The IDEB has developed school goals and measures their progress toward that goal. The ranking provides parents and the public better information about the quality of student and school performance. Families' access to such information also allows them to hold their schools accountable.[111]

17. Individual states have used IDEB to support low-performing schools. For example, in the state of Ceara, IDEB was used to identify the lowest-performing secondary schools. New principals

[103] Ministerio da Educaçao (2005).
[104] Bruns, Evans, and Luque (2011).
[105] Draibe (2004).
[106] By 2010, the ministry had fixed the minimum to be spent on education as 25% of state and municipal revenues and 18% at the federal level; see OECD (2010).
[107] FUNDEF's additional funds allowed poor states in the north and northeast to expand their school provisions and move toward universal elementary education; it also raised elementary school teachers' salaries because 60% of the funds allocated to elementary education were meant for salaries; see OECD (2010).
[108] Bruns, Evans, and Luque (2011).
[109] Sistema de Avaliação da Educação Básica is the National Evaluation System.
[110] A measure that combines test scores and student retention; see OECD (2010). IDEB is superior to current practices in the United States and in many other OECD countries in the quantity, relevance, and quality of the student and school performance information it provides; see Bruns, Evans, and Luque (2011).
[111] Receiving information increases the fraction of parents choosing higher performance schools; see Hastings and Weinstein (2007) cited in Firpo et al. (2016).

are hired in these schools and are required to submit an improvement plan. On a bimonthly basis, superintendents monitor the implementation. In addition, 150 of the best performing schools are given incentive to partner with a low-performing school to help it improve. The highest performers receive additional funds, but one-third of their bonus is contingent on working with a low-performing school.[112]

Studies have found a positive impact of teacher bonus programs (tied to IDEB targets) in the state of Pernambuco on learning outcomes.[113] Under the Pernambuco program, schools are paid a bonus if they achieve at least 50% of their annual performance IDEB targets. The size of the bonus varies depending on progress made. Preliminary results show that average Portuguese (math) scores in 8th and 11th grades rose by 0.44 (0.27) and 0.57 (0.31) of a standard deviation, respectively.[114]

18. With the decentralization process ongoing, and enrollment rising, education quality began to lag.[115] To maintain school accountability, an evaluation framework to assess learning was required to bolster quality. Yet SAEB-based studies showed that the fund did not have a direct positive impact on student performance.[116] FUNDEF created incentives for municipalities to open new schools, increasing enrollment numbers and bringing children from poorer families into the school system.

    In 1990, Brazil lagged far behind middle-income Latin American and Caribbean (LAC) countries and OECD countries on all educational indicators. Less than 40% of children nationally completed the eight grades of primary school, compared with 70% for the LAC region and 95% for the OECD.[117] Rising enrollment had an adverse effect on quality, especially in schools of poor municipalities with a larger contingent of poor students.[118]

    There are a couple of possible explanations why the most recent PISA showed declining scores. The first explanation can be tied to the expansion of secondary education.[119] As of 2015, 71% of 15-year-olds are enrolled in seventh grade or higher as of 2015 (a 15% increase from 2003). Edstats data shows net enrollment in secondary increased from 77% in 2011 to 81% in 2013. Around the same time, the national enrollment rate (NER) in secondary had increased from 74% to 78% between 2006 and 2009, but scores did not drop in 2009. Second, although no link has been suggested, 2015 was the first year that Brazil administered computer-based PISA tests to all participants.[120]

Motivation/drivers in education reform

[112] Bruns, Evans, and Luque (2011).
[113] Bruns and Luque (2014); Ferraz, Finan, and Moreira (2012).
[114] Bruns, Evans, and Luque (2011).
[115] Parandekar, Amorim, and Welsh (2008).
[116] Andrews and Vries (2012).
[117] Bruns, Evans, and Luque (2011).
[118] Bruns, Evans, and Luque (2011).
[119] NeuroMat (2017).
[120] Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira in 2016. "Brazil in Pisa 2015 Executive Summary," Ministério da Educação.

19. The early decade saw a top-down reform team that applied selective coalition building to back the decentralization process. In the 1990s, the greatest pressure to improve the quality of education came from the elites, even though they had no children in public schools.[121] They understood that the future development of Brazil depended upon a well-educated workforce and citizenry.[122] Secondly, social policy was very cohesive, as Cardoso reserved "social ministries," such as health and education, for members of his left-leaning party (PSDB).[123] Thirdly, education interest groups were weak due to the fragmented education system. Fourthly, middle-class families were absent in the public school community because they opted for private schools.[124]

    Partnership strategies have become stronger instruments throughout the Brazilian education system reform. Despite earlier low social actor engagement, parent-teacher associations (PTA) gained some prominence.[125] In this context, FUNDEF funding, for instance, helped shift autonomy to subnational governments and school directors to pursue programs such as PMDE.

20. The second wave of reforms in mid-2000 came with the shift toward not only consolidating access but also improving quality**.** The second wave meant moving away from a fragmented policy terrain toward a more unified education strategy. In 2006/2007, FUNDEB extended preschool and secondary coverage and also explicitly guaranteed minimum levels of per-capita funding for enrollment in education programs for indigenous and *quilombo* communities and youth and adult education. The process depended on public partnership movements to maintain accountability. For instance, the Compromisso Todos pela Educação initiative (2005) reinforced efforts to adopt better IDEB levels using good practices.[126] This enabling environment helped reinforce the foundational pillars for improved learning.

Reform process

21. Inadvertent power shifts surfaced in the reform process. Since FUNDEF gave low incentives for state actors to support such level of decentralization, reformers moved swiftly to pass approval. The reformers' use of political clout, i.e., the president, first lady, and political bargaining power of the Education Ministry, were a leverage to gain congressional approval. Their strategy also hinged upon strong incentives for municipal authorities to endorse primary education. The school management program (PMDE), for instance, saw high teacher and principal endorsement, having the dual goal of also engaging PTAs. The PMDE stipulated that funds be delivered directly to municipal authorities or school directors until the required school board was setup (PMDE covered 90% of schools by 1998).

Role of coalitions

[121] The former president of the National Institute for Educational Studies and Research; see Draibe (2004).
[122] OECD (2010).
[123] Draibe (2004).
[124] Draibe (2004).
[125] Brazil Ministry of Education (2015).
[126] Parandekar, Amorim, and Welsh (2008).

22. Coalition tactics became crucial during the reform phases. Yet to pass the PMDE meant federal power relinquished key bargaining power to states and municipalities.[127] While aligning the ministry to gain state governor and mayor support, they forged coalitions with local counterparts at three levels: states, municipalities, and school units. PMDE was approved by ministerial decree, not congressional legislation. It meant that automatic criteria for resource allocation eliminated stakeholder negotiation about sums.

23. Brazil's reforms have shown major progress over the last few decades in redressing deep inequities. Changes were rooted in the urge to develop the economy more equitably. Education finance is a clear example, where education officials established highly innovative tools backed by political dedication. The strategy of the three administrations[128] was consistent in its approach to reaffirm the efforts made to mobilize resources for education. One of the most important examples of policy continuity between the Cardoso and Lula da Silva administrations was when they reauthorization and expansion of FUNDEF in 2007 as FUNDEB. Another example is that FUNDEF even included "a sunset clause" after eight years.

24. The FUNDEB revenues have driven education spending upward and equitably distributes to basic and secondary education.[129] FUNDEB was even transformed from a formula based on population density (i.e., biased to large cities) to a student-based funding formula (based on minimum per-pupil allocations).[130] In other words, regardless of where students attend schooling, they have enough learning resources (for desks, electricity, water, books, pencils, and books).[131] This alone does not guarantee improved learning, and the following paragraphs explain how both IDEB and Prova Brasil influence learning.

Role of information, innovation, and agility

25. Innovation and partnership-building have spurred the reform process toward better education. The accountability made through decentralization mechanisms and IDEB targets have endured over time. The swift coalition strategies that the ministry had exerted to approve the fund helped manifest accountability and social controls locally. This helped the decentralization of education provision and helped expand its reach.

26. Brazil aspires to world-class learning standards, and performance targets are already above expectations in the country's vision to meet average PISA scores by 2021. Brazil has also institutionalized IDEB. The groundbreaking index tool is a major achievement, and it links

[127] Draibe (2004).
[128] The latter two being the Workers Party.
[129] Bruns, Evans, and Luque (2011).
[130] Yearly FUNDEF tops up education resources for around 6 of Brazil's 26 states. For instance, the states of Acre and Ceará were given the necessary resources to improve their education systems.
[131] FUNDEF also mandates that 60% of the total per-student allocation be spent on teacher salaries and 40% on other operating costs. The impact of the mandate in its first several years was a 70% increase in average teacher salaries in poorer municipalities in the northeast and north; see Gordon and Vegas (2005) as cited in Bruns, Evans, and Luque (2011).

examination results (ProvaBrasil scores and pass rates) to performance in learning assessments. Between 2005 and 2009, national learning rose to outperform all target scores across the school cycles. Performance has risen from 3.8 to 4.6 for primary schools (grades 1–4) between 2005 and 2009, above the 4.2 target; from 3.5 to 4.0 in intermediate schools (grades 5–8), outperforming the target of 3.7; and from 3.4 to 3.6 in high school (grades 9–11), outperforming the target of 3.5.[132]

27. Participating and planning motivate better learning. IDEB was backed by the Compromisso Initiative, with a Plan for Coordinated Action, or PAR, an acronym that means "partner" in Portuguese. Public funding transfers hinge upon a prerequisite school planning with i) the Political Pedagogical Project and ii) the School Development Plan (PDE-Escola). Plans must be developed with the participation of teachers, and parents must be notified of the plans' execution.

    The index results have also created public pressure that incentivizes schools to improve based on its passed performance.[133] Ms. Pilar, secretary of Basic Education in 2010, recounted a visit to a school in a difficult area of Rio de Janeiro where 1,000 parents and community members were celebrating the release of the IDEB scores. Schools that show great progress are given more autonomy, while schools that remain low performers benefit from more help. In 2008, the Ministry of Education prioritized work with the 1,827 lowest-performing municipalities, providing resources and technology.

28. Brazilian municipalities are electorally accountable for the municipal school system's poor performance. Firpo, Pieri, and Souza (2017) show the link between school quality and electoral accountability of incumbent mayors (elected every four years).[134] From 2005 to 2007, gains in school quality in municipal schools increased chances of reelection among eligible (for a second term) mayors in the 2008 election.[135] "This effect is even greater in localities with lower per-capita income (higher demand for public education) and those where the fraction of children at school age is larger." Local governments (in relation to mayors) have recently also adopted performance incentives for public school principals.

29. Healthy rewards and bonus incentives for improved schools catalyze motivation for better quality education. Simultaneously, the index also discourages schools from holding children back to boost learning scores. "The use of IDEB has changed the relationship between the ministry and municipalities and states," explained ministry executive secretary, Mr. Paim. "States have to diagnose the problems in low-performing schools and develop an improvement plan to send to the ministry."[136]

[132] OECD (2010).
[133] OECD (2010).
[134] Firpo, Pieri, and Souza (2017).
[135] A one-point increase in a 0 to 10 scale index from 2005 to 2007 increased the probability of mayoral reelection by around 5% points.
[136] Avoiding incentives for grade retention is crucial in Brazil, as the average repetition rates in primary school are approximately 20%, the highest in Latin America; see OECD (2010).

30. Brazil is motivated to assert itself as a competitive player on a global scale, but it still grapples with issues of access, quality, and equity. System coherence and alignment are essential. Chile's approach of first establishing the standards and then aligning tools is effective, but Brazil's federal system makes system coherence more difficult. Brazil demands high standards and tackles changing certain aspects of the system rather than the whole system at once. Given this context, Minister Fernando Haddad's proposal of teachers having to take an exam before entering the classroom, and Sao Paolo Secretary Renato Souza's (former Minister of Education) teacher career plan, are welcome efforts put forward to improve qualified teaching staff.

Lessons learned

31. Brazils' pivotal political actors pushed the reform agenda ahead. Vision and leadership were major success factors in improving the education system, with reformers bringing about a more equitable education framework. Brazil's education system is diverse where autonomy permits flexible education legislation plus school system designs. Coalitions with local counterparts have also become crucial in municipalizing funds, as some political bargaining inadvertently shifted agency from central to regional spheres due to decentralization.

32. In turn, reform innovation helped encourage better school systems. Information drove new targets spurred on by more engaged citizens. The IDEB has become a widely known tool that has catalyzed schools and its communities to demand better learning and to outperform their past accomplishments. Not only was information a motivator in its own right, but it also had a political dimension.

33. Reforms have played major roles in accountability. The devolution of FUNDEB and the impacts of IDEB data are key. The decentralized funding has empowered local government, especially mayors, and communities have also become informed, further engaging its officials. The role of the IDEB performance information has incited quality and has helped push demand for an improved school system performance. Gains in quality schooling meant a higher probability of reelection among eligible mayors in the 2008 election.

34. Coverage rose, yet typically at a cost to quality. Brazil's administrations have used effective reforms and have set up a framework to help expand basic quality education. The role of a unified strategy of information gathering and disseminating has also helped hold school bodies accountable to improve learning and quality. High standards and completion rates should remain prime policy areas throughout further reforms, and foundations for better learning, such as including preschool, are key. Early childhood development (ECD) studies, for instance, show that students enrolled from preschool perform better than students who enroll at the primary school level.[137]

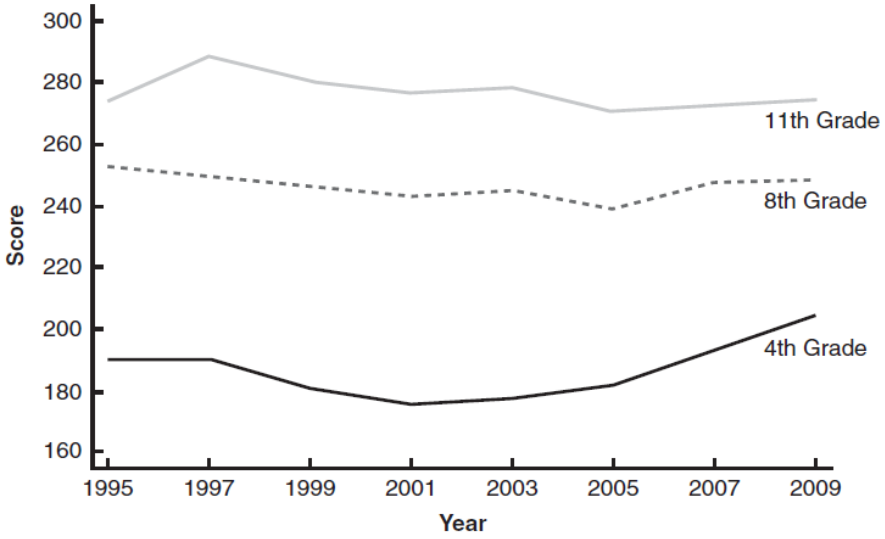
35. While inequality, poverty remains a key explanatory factor for school performance levels. The gradual increase in the average 15-year-old student's schooling years remains a central

[137] Parandekar, Amorim, and Welsh (2008).

determinant in explaining positive learning outcomes, especially in math.[138] Inequality has declined in a country considered historically, highly unequal, an 8-point drop over two decades (the Gini index was 60 in 1993, 58 in 2003, and 52 in 2015).[139] However, the impact of poverty on school performance remains a challenge.[140]

36. Reforms that aim to improve learning rely on strong strategies and good partnerships. Besides political coalition building, partnerships also began contributing to healthy incentivizes, and these led local officials and principals to fulfill demands for better learning in schools (the Compromisso Initiative backed IDEB results).

Figure A2: Math Proficiency on SAEB/Prova Brazil, 1995–2009



Source: Bruns, Evans, and Luque (2012).
Notes: National SAEB (1995–2005); SAEB/Prova Brasil (2007–2009). Public schools only. For 4th and 8th grades, state and municipal schools. For 11th grade, state schools.

[138] The Emilio Garrastazu Medici Administration ensured that primary and secondary education reforms extended compulsory basic education from four to eight years; see Draibe (2004).
[139] World Bank (2017).
[140] Andrews and Vries (2012).

Country case 2: Passing Chile's pay-for-performance reforms (1995–2005)

Introduction

1. Reforms tied to evaluating and improving teacher standards and linking teachers' pay to performances are typically short lived and achieve limited success.[141] Nevertheless, Chile's experience in negotiating and implementing teacher performance incentive reforms during the mid-1990s to early 2000s is an exception. For example, during this period, the Ministry of Education (MINEDUC) passed several reforms directly related to teacher performance. Using an inclusive strategy and a phased approach to introduce technically sound reforms, Chile implemented school-based bonus pay and individual pay incentive programs tied to performance assessment, an implementation still in practice.[142]

2. Initially, the union resisted any proposal to introduce collective performance incentives.[143] However, they eventually supported the program and codesigned subsequent policies on individual-level evaluations. By the late 2000s, 15%–25% of teacher salary was linked to collective and individual-level performance-based incentives.[144] This period of sequenced teacher incentive reforms also coincided with improved learning outcomes. These reforms, implemented more than ten years ago, have remained relatively stable over time.

3. The Chilean experience stands out not only as an exceptional case of success in implementing pay-for-performance type reforms but also provides an example of reforms that have addressed technical and political misalignments at the system level. This case study highlights this key feature of the Chilean education reforms and their contributing factors. The purpose of the case study is not to promote Chile's education reforms as best practice but to identify the strategies and factors behind successful reform episodes.

Data/evidence: Trends in learning outcomes and definition of reform episode

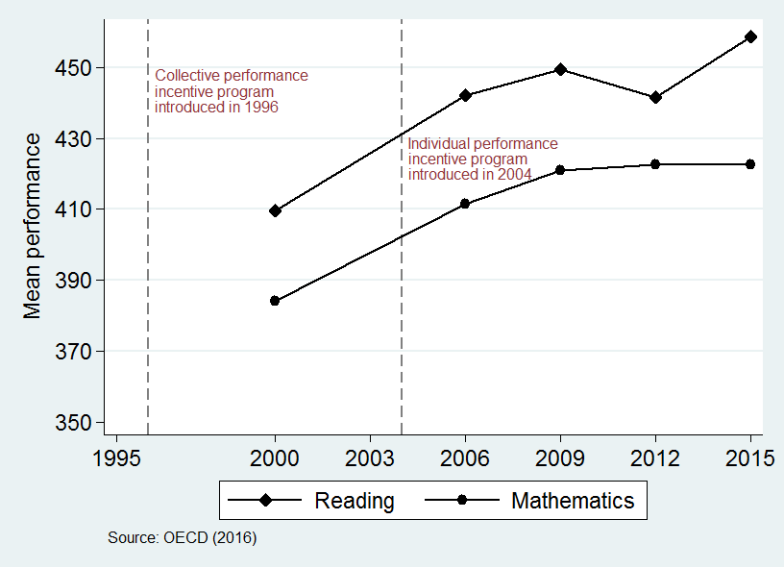
4. Chile experienced a significant and sustained improvement in learning outcomes from 2000 to 2015. As Figure A3 shows, the biggest jump in mean performance in the PISA tests[145] in math and reading occurred between 2000 and 2006. These improvements came after collective and individual performance incentive programs were introduced in 1996 and 2004. Results from the national assessment system, Sistema de Medición de la Calidad de la Educación (SIMCE), also showed improvements at the primary level.[146]

---

[141] Some examples are the ongoing protests in Oaxaca, Mexico in response to a proposal for mandatory teacher evaluations and protests by unions in Peru and Ecuador.
[142] The recently passed National Teacher Policy in 2016 expands these policies to teachers in voucher schools and institutes further changes to teacher recruitment and development.
[143] Mizala and Schneider (2014).
[144] Mizala and Schneider (2014).
[145] A nationally representative sample of 15-year-old students participate in the PISA tests.
[146] Scores among fourth-grade students increased by 11 points in math and by 17 points in reading between 1999 and 2012. See https://s3.amazonaws.com/archivos.agenciaeducacion.cl/informes-resultados–2012/Conferencia+Nacional+de+Resultados+SIMCE+2012++II+medio+y+4+basico.pdf.

5. Improvement in learning outcomes coincided with expanding coverage. While gross enrollment at the secondary level was 87% in 2000, by 2015 Chile had achieved universal enrollment.[147] Per-student spending increased by 60% (73%) at the secondary (primary) level over the 15-year period. Secondary enrollment jumped by over 10 percentage points between 2000 and 2006 from 87% to 98%. Spending on education also increased between 2000 and 2015. The sharpest increases in education spending, however, occurred post-2006, after the biggest improvements in learning had been realized. Real per-student spending at the secondary level actually declined between 2000 and 2006 by 7%.[148]

Figure A3: Mean scores in reading and math over time



Source: OECD (2016).

Broader context: Overall political and economic context

6. After two decades of dictatorship under Augusto Pinochet, during which public expenditure on social services including education had been low, the early 1990s marked a return to democracy and to increased attention to public education. The Pinochet regime had decentralized education and encouraged the private sector's role in education to foster competition.[149] The regime had also stripped public school teachers of civil servant status and associated benefits, resulting in a demotion in status and economic conditions.[150]

[147] UIS (2017).
[148] Spending per secondary student (in constant USD) was $1,445 in 2000, while in 2006 it dropped to $1,340 (UIS, 2017).
[149] The basic system structure of the 1980s still remains. Schools can broadly be categorized into three groups: municipal schools (receive public financing and are managed by municipal authorities), voucher schools (private schools receiving government subsidies), and private schools (do not receive any government support); see Pont et al. (2013).
[150] The starting salary for a teacher in 1990 fell below the poverty line; see Crouch (2005).

Chile's primary teachers' union, the Colegio Docente (hereafter referred to as the Colegio) convinced the post-democratic Aylwin Administration to officially restore their status through the Teacher Statute in 1991. This statute not only improved teachers' conditions but also set the precedent for future reform negotiations with the union such as collective and individual performance incentives.

7. Relative political stability in the two decades after the military regime ended in 1990 may have also contributed to the reforms' success. The first three consecutive governments during the crucial transition period to democracy belonged to the same coalition of parties, the Concertacion. In addition, Ricardo Lagos, the education minister during the first post-democratic administration, went on to become president from 2000 to 2006, during which the government introduced teacher performance review and individual-level incentives.[151] Moreover, the Colegio belonged to the same coalition that upended the Pinochet regime, which might have also helped foster dialog over typically contentious reforms.

8. High economic growth and stability[152] continued during the post-democratic period, which may have created the political space for introducing reforms and improved learning. In the first decade after democracy, per-capita GDP grew annually by 5%, on average. For the next five years, the economy continued to grow at an average annual rate of approximately 3%.

Education reform program

9. In 1996, the MINEDUC, under the Frei Administration, introduced a program that tied bonus pay for teachers to school-based performance through the Sistema Nacional de Evaluacion del Desempeno de los Establecimientos Educacionales (SNED). In addition to assessing schools' academic performance, SNED introduced the first collective incentive for teachers and paved the way for individual-level incentives about a decade later.

   Ninety percent of the incentive amount was distributed among teachers, while the use of the remaining 10% fell on the discretion of the school principal (this distribution has remained the same). Municipal and voucher schools at each level were categorized into groups based on geographical region and students' socioeconomic characteristics.

   Schools were (and are to this day) evaluated based on five main variables that included absolute and relative performance in the national-level student assessment, SIMCE; outcomes measuring completion and equality of opportunities; integration of teachers and parents; improvements in staffing conditions and initiatives on pedagogical activities; school development plans; and creation of teacher councils. (Refer to Annex Table A1 for a full description of the components and their weights.) Schools were evaluated every two years, and the top performing schools[153] within each group were awarded a bonus.

[151] Mizala and Schneider (2014).
[152] Unlike its neighbors, such as Mexico and Argentina, Chile avoided economic crises in the 1990s.
[153] The top performing schools represented up to 25% of enrollments in each group.

10. In addition to the SNED program, MINEDUC instituted a package of initiatives that fostered teachers' professional development and aimed to recruit strong candidates. These initiatives included a study tours program for about 3,000 beneficiaries to provide exposure to best teaching practices, in-service training programs, and a competitive funding provision for 17 universities to align teacher education with requirements of the broader education reform package.[154] The administration also provided scholarships to 300 students per year to attract good students into the teaching profession.

11. The early 2000s witnessed a series of even more ambitious reforms that linked teacher pay to individual evaluations in spite of uncharacteristically long teacher strikes over base salaries in the late 1990s.[155] Following negotiations with the Colegio, in 2002, MINEDUC introduced a voluntary individual-level incentive for teachers in municipal and voucher schools. The Asignacion de Excelencia Pedagogica (AEP) allowed the evaluation of teachers' classroom performance and content knowledge on a voluntary basis. Those who passed the evaluation were eligible to receive bonuses for the following ten years. Moreover, teachers deemed high performing in the AEP had the chance to propose mentoring of other teachers and were eligible to receive financial incentives to do so.[156]

12. The most significant feat during this period was when the administration adopted mandatory teacher evaluations and linked individual-level incentives for municipal school teachers. After negotiations and consultations with Colegio members over four years, the individual teacher evaluation system, known was Docentemas, started in 2004.[157]

    Teachers would be evaluated every four years on their competency as well as qualitative assessments based on classroom practice, self-assessments, and peer evaluations. They were ranked into four grades of "outstanding," "competent," "basic," or "unsatisfactory." Those ranked in the top two performance categories were rewarded by having priority in any promotion opportunities. Furthermore, upon passing a voluntary test on curricular and pedagogical knowledge, top performing teachers received individual bonuses under the Asignacion Variable por Desempeno Individual (AVDI) program on a graduated scale based on whether the teacher was "outstanding" or "competent."

13. The Chilean teacher reforms from the 1990s to early 2000s represent good examples of well-designed policies that resolved technical and political issues in aligning the system toward improving learning. Rigorous evaluations of the SNED program show its efficacy in improving learning outcomes. Contreras and Rau show a positive average treatment effect among all

[154] For example, the set of education reforms passed in 1996 included a new curriculum framework that emphasized competencies, such as critical thinking, abstract reasoning, problem-solving, communication etc., as well as other learning objectives such as moral values; see Delannoy (2000).
[155] For example, Mizala and Schneider (2014) mention a two-week strike by the Colegio in 1996 and another month-long strike in 1998. However, they find that the strikes were mostly over base pay rather than against performance-based incentive programs.
[156] Crouch (2005).
[157] Crouch (2005); Mizala and Schneider (2014); Pont et al. (2013).

schools that were eligible for the program.[158] They find that SIMCE scores at the primary and secondary levels in language and math increased by 0.16 to 0.25 standard deviations.

14. The design of the SNED program incorporated a holistic approach that not only resolved potential technical misalignments but also reduced chances of political opposition. While significant weight is placed on measurable learning outcomes, the program also takes into account innovations undertaken by the school and working conditions that foster teachers' professional development. Including such measures is likely to direct schools toward improving performance beyond simply scoring well on the SIMCE tests.

    To ensure fairness in chances of winning the bonus, schools are compared within homogenous groups based on region, a school vulnerability index, and other measures of socioeconomic background of students, and location type.[159] Comparing similar schools addresses the possibility of discriminating against teachers in poorer schools that may not be predisposed to performing as well as wealthier schools. To further ensure impartiality, MINEDUC contracts out implementation to a department at the University of Chile.[160]

15. As in the case of the SNED, the individual teacher evaluation system (Docentemas) was carefully designed to address potential misalignments. MINEDUC developed the Good Teaching Framework in 2003 to use as standards to evaluate teachers. The evaluation process goes beyond measuring knowledge of curricular content and pedagogical skills. Instead, teachers are assessed using qualitative measures based on a teaching portfolio, self-assessment, peer evaluations, and a third-party reference.

    While peer evaluators need accreditation and familiarity with the evaluated teachers' subject area, location, and socioeconomic conditions, they may not be from the same school. This feature helps avoid the pitfall of encouraging disincentives to impartially rate peers. Moreover, independent third-party university research centers implement the whole process, from producing instruments to selecting evaluators and information processing.[161] This feature also lends objectivity to the evaluation process.

16. The Colegio's acceptance of individual-level evaluations and incentives is explained by the fact that while the performance evaluations are mandatory for municipal teachers, qualifying teachers can choose to undergo further testing on curricular and pedagogical knowledge and can earn bonuses. On the other hand, making bonus pay conditional on passing competency tests helps prevent the individual incentive program from draining the public education budget. It is also noteworthy that neither the AEP nor the AVDI are directly linked to student learning outcomes, which usually tend to be politically contentious.

17. Adopting a "carrots rather than sticks" approach to raising teacher standards could have made the Docentemas program less unpopular. For example, teachers placed in the bottom ranks

[158] Contreras and Rau (2012).
[159] Pont et al. (2013).
[160] The Center for Applied Economics of the Industrial Engineering Department at the University of Chile develops the instruments and computes the school-level indices in the SNED; see Pont et al. (2013).
[161] Pont et al. (2013); Avalos and Assael (2006); Bruns and Luque (2014).

have opportunities to improve through a professional improvement plan.[162] Moreover, those deemed "unsatisfactory" must go through a follow-up evaluation after one to two years, depending on initial evaluation ratings. It is only after "unsatisfactory" ratings in three consecutive annual evaluations that teachers are dismissed.[163]

18. Since the period from 1995 to 2005 was characterized by increases in teacher salary, it is tempting to view these salary increases as compensation to teachers and the Colegio to buy-in support for the reforms. However, Mizala and Schneider (2014) note that while salary increases may have helped generate support and establish credibility for the government, increases in teacher salaries did not only occur in periods around reform negotiations. Instead, Mizala and Schneider note that increasing teacher salaries also served the purpose of attracting qualified individuals into the teaching workforce in the long run.[164]

Motivations/drivers of education reform

19. While Chile had already achieved relatively high rates of access,[165] education quality in the 1990s was characterized by substandard outcomes in learning and equity issues in education. Assessments conducted at the beginning of the 1990s revealed stark realities on student performance and inequity in the Chilean education system. For example, 40% of poor fourth-grade students scored below basic competency levels in reading. In addition, the repetition rate was nearly 8% in basic education, contributing to an average completion time of more than ten years for a basic education cycle of eight years.[166]

    Galvanized by such statistics, the post-democratic administrations embarked on a series of reforms that included programs designed to improve learning and hold schools and educators accountable. Having high levels of education coverage by 1990 may have also made it easier to shift attention to improving quality.

20. Strong technical and visionary leadership on the side of the administration, as well as the Colegio, also played a role in how the reforms were designed. Solid technical leadership in the education and finance ministries allowed for collaborations between the two to design sound policies. Although the Colegio enjoyed monopoly status, its involvement in education policy went beyond typical union concerns. For example, the Colegio prided itself on its high technical and intellectual capacity and was open to consultations from external advisors.[167]

Reform process

Implementation

[162] Bruns and Luque (2014).
[163] Avalos and Assael (2006); Mizala and Schneider (2014).
[164] Mizala and Schneider (2014).
[165] As of 1990, enrollment at the primary level was universal. Gross enrollment in secondary was 77%; see UNESCO (2017).
[166] Delannoy (2000).
[167] Crouch (2005).

21. The collective incentive program was implemented without much controversy since consensus had already been achieved during the design stage. As per the original design, the administration evaluated schools every two years. In the first five rounds, the program benefited 20% of schools and 28% of teachers. After the initial rounds, the program itself was evaluated to guide implementation in subsequent rounds.[168] In 2004, the size of the bonus doubled, and this bonus amounted to approximately half of a monthly salary per year for each teacher in the recipient schools.[169] The program has been codified into law and remains one of the longest lasting teacher performance incentive programs in Chile.

22. Implementing the individual incentive programs faced some challenges. Assessments under the voluntary individual incentives (AEP) had a low pass rate. Mandatory individual evaluations and incentives (AVDI) faced initial resistance from a minority of rank-and-file union members. By the end of the second year of introduction, over 5,000 teachers had refused to participate in evaluations. However, findings from national surveys reveal that a larger share of teachers supported individual performance incentives. Surveys also show that opposition among teachers to individual incentives fell over time.[170] However, by 2010, about one-fifth of municipal school teachers received bonuses through the AVDI.[171]

    Support for performance incentives was higher among school principals, with 84% of principals reportedly expressing support in 2006. Moreover, support was higher among principals from schools that had received SNED bonuses, suggesting that introducing SNED may have helped build acceptance for performance evaluations in general.

Information and knowledge

23. Strategic use of available information has been a key driver of Chile's teacher reform experience at various stages. Information on schools were available to parents through the Programa de Evaluacion del Rendimiento (PER) as early as the Pinochet regime. While SIMCE results were published only at an aggregated level before 1995, available information showed the poor status of learning and gaps in equity between private and municipal schools. Crouch suggests that the availability of such information, particularly in light of education spending increases, might have played a role in creating an urgency for implementing reforms to improve quality.[172] Moreover, Chile's participation in international assessments in the late 1990s and early 2000s provided valuable information, revealing Chileans' low performance relative to other countries.[173]

24. To help implement SNED, MINEDUC undertook steps to align reporting of the existing SIMCE data with the program's requirements. Although the SIMCE had been implemented

[168] Delannoy (2000).
[169] Mizala and Schneider (2014).
[170] Surveys conducted by Chile's Center for Education Research and Development show that the proportion of municipal school teachers who opposed individual incentives decreased from 44% in 2004 to 36% in 2006; see Mizala and Schneider (2014).
[171] Mizala and Schneider (2014).
[172] Crouch (2005).
[173] Chile participated in the TIMSS in 1999 and in the PISA in 2000.

since 1988 to encourage school competition, school-level results were not reported until the Frei Administration changed the reporting policy in 1995. The administration also adopted a technique to equate scores across different test years to ensure the comparability needed to know whether or not a school experienced growth in student outcomes.

25. In addition to information on achievement, the breadth of information collected through the SIMCE's household questionnaire enabled MINEDUC to create a sophisticated design for the incentive programs. For example, under SNED, schools compete with other schools with similar demographic and socioeconomic characteristics such as average household income and parental schooling.

Coalitions and incentives

26. Several factors lie behind Chile's atypical success in passing and implementing pay-for-performance type reforms. While some factors are related to pre-existing political conditions, capacity, and policies favorable to the reforms, discretionary strategies also provided incentives to different actors to support the reforms. For example, policy makers were cautious to not use the term "reform" for fear of the negative connotation associated with it at the time. Instead, they referred to the proposed policies as "policy changes." Such care in how these reform proposals were communicated shows the expediency with which the administrations and policy makers handled the reform process.

27. Coalitions formed before and during the reform process helped pave the way for successful reforms. Centralized salary negotiations under the 1991 Teacher Statute set the stage for regular negotiations and dialog with the Colegio instead of confrontation on other policy changes.[174] This step and subsequent salary increases helped establish the credibility of the Concertacion government among the Colegio. The continuity in administration by different parties of the Concertacion during the following three terms helped continuity and also helped policy makers build on previous teacher reforms. Efforts in the early 1990s to encourage teachers' development through support networks may also have placed the government in the Colegio's good books.[175] Such steps may have also made the Colegio less opposed to proposed reforms.

28. Through most of the reform process, the government adopted a consultative strategy where stakeholders, including the Colegio, actively participated in various stages, from policy design to implementation. For example, in 1994, the Frei Administration established a commission composed of actors ranging from educators to academics to the Catholic Church in order to review needs in education. SNED was proposed based on this commission's recommendations. Colegio representatives codesigned the SNED and the individual-level incentive programs.[176] In addition, a committee comprising representatives from the Colegio, Chilean Association of Municipalities, MINEDUC, and academics gave advice on the teacher evaluation process.[177]

[174] Mizala and Schneider (2014).
[175] For example, the Mejoramiento de la Equidad y de la Calidad de la Educacion projects in partnership with the World Bank introduced processes to build professional networks of teachers; see Delannoy (2000).
[176] Crouch (2005).
[177] Avalos and Assael (2006).

Although the Colegio was initially opposed to SNED, endorsement by other political parties and actors discouraged fierce opposition. Moreover, factors such as salary increases and job creation from the expansion of secondary schools helped buy the union's support.

29. Moreover, officials across sectors and agencies participated in various stages of the reform process by helping to design and implement the reforms. For example, the key designers of the SNED were policy makers from the Ministry of Finance who later joined the top-level administration within the Ministry of Education at the time of implementing reforms.

Innovation and agility

30. Successive administrations during the reforms employed innovative approaches to ensure teacher reforms were accepted. The teacher evaluation system and pay-for-performance program were not conceived in a vacuum; proposals for some form of review and teacher accountability dated back to the early 1990s.[183] Yet despite available technical capacity, the Concertacion government adopted a principle of gradual implementation versus a single sweeping reform package. For example, a 1999 committee report on the potential design of teacher evaluations recommended a gradual approach and emphasized the need for developing an "evaluation culture."[184]

31. Sequencing reforms from collective to mandatory individual-level incentive programs was key for the political success of the reforms. Furthermore, before introducing mandatory individual-level evaluations and incentive programs in 2004, the Lagos Administration tested the waters by introducing voluntary individual-level evaluations (the AEP) in 2002. Having set a precedent for individual evaluations, the administration developed the mandatory teacher performance evaluation system for municipal school teachers.

32. Another key strategy of the Concertacion government relied on using broader education reforms as entry points for introducing performance-based incentives rather than proposing the incentive reforms in isolation. For example, SNED was introduced along with the Full School Day reform package as part of a teacher professionalism pillar.[178]

    The introduction of the reforms were also generally preceded by increases in base teacher salaries, which may have made teachers more amenable to some form of performance-based incentive programs. For example, the introduction of the individual-level incentive was preceded by guaranteed salary increases for a three-year period. In addition, the reform proposal was also accompanied by an expansion in coverage of schools eligible for SNED and an increase in the associated incentive amount.[179]

Lessons learned

[183] The 1991 Teacher Statute, for example, included a provision that required all public sector employees have a performance evaluation; see Avalos and Assael (2006); Mizala and Schneider (2014).
[184] Avalos and Assael (2006).
[178] Delannoy (2000).
[179] Mizala and Schneider (2014).

33. Chile's success in introducing pay-for-performance type reforms that have sustained over time has been atypical. While it would be unreasonable to expect all countries to experience favorable conditions, such as continuity in administration or ample resources for continued increases in teacher salary, there are some lessons that can be drawn from the Chilean experience.

34. Policy makers could seize appropriate windows of opportunity for reforms to improve quality. For example, Crouch suggests that collective or individual-level incentives may be easier to negotiate when average base teacher salaries are increasing.[180] Authorities could seize opportunities of salary recovery to introduce teacher-related quality reforms. On a related note, including typically contentious quality reforms along with less contested policies may be useful, as Chile did when introducing individual-level incentives.

35. Opting for incremental changes to policy over time, rather than a drastic change, can be a more prudent way to gain acceptance and pass reforms. Chile's experience has shown that sequencing reforms beginning with those less likely to be opposed helps increase the acceptance of less popular reforms. Furthermore, such an approach is more likely to guarantee that the administration implements meaningful reforms rather than launch a drastic policy change that may be rejected in the beginning. In the case of teacher accountability, starting with collective incentives can help build an evaluation culture, making it easier to implement individual incentives that could be directly tied to learning.

36. Developing and using learning information appropriately can help build demand for quality reforms and can allow for well-designed reforms. Given the role that key information on learning in the 1990s played in motivating quality reforms in Chile, it would be wise to develop effective learning information systems to help make the case for quality-enhancing reforms. By adding features to the SIMCE test, such as cross-time comparability and using existing information on socioeconomic characteristics, MINEDUC developed a technically sophisticated incentive program. Furthermore, the decision to not link individual teacher incentives with SIMCE test scores prevented SIMCE from turning into an unpopular high-stakes test rather than a diagnostic tool.

37. A consultative approach to design reforms involving various actors in education and across administrative sectors is likely to build support at the time of implementation. In designing the collective incentive program, MINEDUC consulted not just with teachers but also with municipal authorities and officials in relevant line ministries such as the Ministry of Finance. This approach helped the administration smoothly implement the SNED. Moreover, granting advisory roles to municipal authorities for the teacher evaluation process may also have incentivized proper implementation in a decentralized framework.

Table A2: SNED components and weights

Variables	Indicators	Weight, 1996–97	Weight, 1998–99

[180] Crouch (2005).

Effectiveness	SIMCE scores in math and Spanish	40%	37%
Value added	Average SIMCE score gain from previous year	30%	28%
Capacity for initiative	Creation of teacher councils Participation in microcenter[181] meetings Pedagogical activities Student council School development plan Teacher workshops	6%	6%
Improvements in working conditions	Full staffing Replacement of absent teachers	2%	2%
Equality of opportunities	Student retention rate Student graduation rate Differential groupings Integration projects Absence of discriminatory practices	12%	22%
Integration of teachers and parents	Acceptance of educational work by parents, guardians and students, and creation of parents' centers	10%	5%

Source: Delannoy (2000).
Notes: SIMCE stands for Sistema de Medición de la Calidad de la Educación and SNED for Sistema Nacional de Evaluacion del Desempeno de los Establecimientos Educacionales.

Country case 3: England's 2007 National Literacy Strategy

Introduction

1. Over the last two decades, England has improved literacy and numeracy skills for primary school-aged children. Between 1995 and 2015, the proportion of fourth-grade students reaching the intermediate benchmark in the TIMSS math assessment increased from 54% to 80%. Few countries that have participated in TIMSS can claim the types of success England has managed to achieve within a short amount of time.[182]

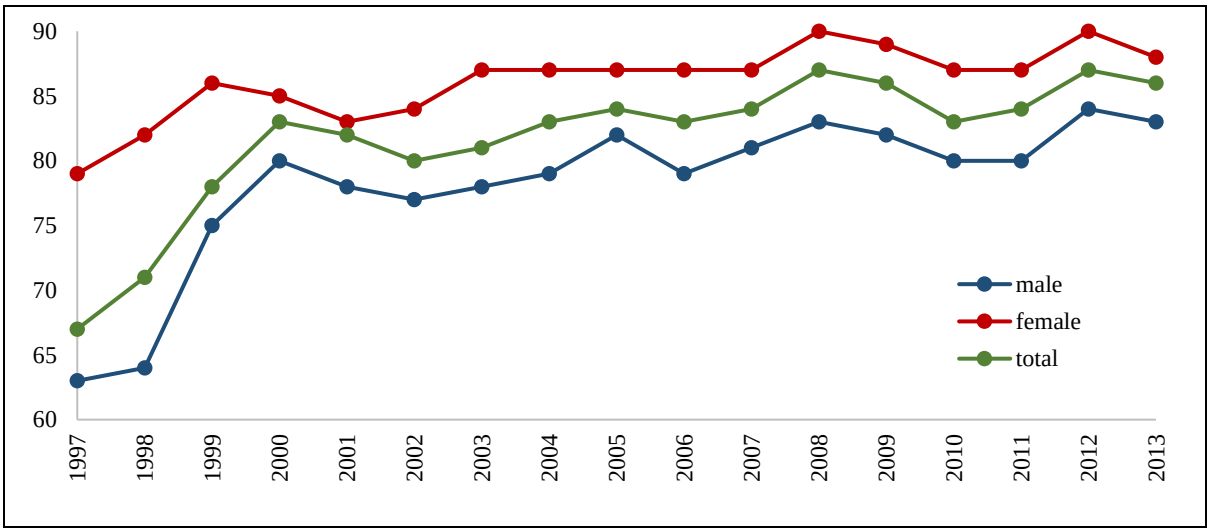
2. This case study offers an account of how the government used technical and political factors to improve learning outcomes through a combination of improving information and metrics, building coalitions and strengthening incentives, and encouraging innovation and agility. The New Labour party, who had just assumed power in 1997, introduced the reform package. Called the National Literacy Strategy, the reform focused on providing a new framework to improve the teaching of literacy and numeracy in primary schools. The program relied on evidence-based strategies and established clear targets (based on student achievement data) for learning outcomes, both nationally and at the school level. This case study highlights the political context in which these reforms were brought about.

Data/evidence: Trends in learning outcomes and definition of reform episode/period

[181] Microcenters host professional networks of schools to discuss pedagogical issues and share solutions. The meetings receive technical support from MINEDUC; see Delannoy (2000).
[182] Mullis et al. (2016).

3. There is substantial evidence, based on impact evaluation work, about the effects of the National Literacy Strategy.[183] A summary of multiple proficiency tests confirm modest and sustained rises in literacy, writing, and math.[184] The proportion of English primary school students in the fourth grade who reached the intermediate benchmark in the TIMSS math assessment increased from 54% in 1995 to 80% in 2015.[185] In 1996, 57% of students achieved Level 4[186] in Key Stage 2 English; by 2002, this had increased to 75%. Between 1997 and 2002, the percent of students reading at Level 4 in Stage 2 increased from 67% to 80%.[187] Despite these successes, the program has continued to evolve and has introduced more focused support to reach the remaining disadvantaged learners.[188]

Figure A4: Improving primary school literacy in England: Percentage of children at the end of primary school (11 years old) reaching Level 4 proficiency on reading by sex, 1997–2013



Source: Cassen, McNally, and Vignoles (2015).

Broader context: Overall political and economic context

4. In 1997, New Labour assumed power from the Conservatives. They came into power with a strong majority as well as a mandate to implement social sector reform. They brought about

[183] Cassen, McNally, and Vignoles (2015); Tanner et al. (2010); Stannard and Huxford (2007).
[184] Tymms (2004).
[185] Mullis et al. (2016).
[186] Children at Level 4 are expected to show understanding of significant ideas, themes, events, and characters to begin using inference and deduction. They can refer to the text when explaining their views and locate and use ideas and information.
[187] Machin and McNally (2004).
[188] Stannard and Huxford (2007).

reforms during a time when England had strong economic growth and per-capita incomes were rising.

5. During the elections, the issue of poor education outcomes had been a concern for both major political parties. The National Literacy Project had been first introduced under the Conservative government. The project formed the National Literacy Strategy, which had been initially introduced at a limited sale within a small number of Local Education Authorities in 1996.[189] Before the 1997 election, an outline of the National Literacy Strategy was launched, and a summary of the report was shared with each primary school across the country.[190] There was sustained pressure to improve children's learning across the country. Both parties recognized that reforms needed to happen, and improved monitoring and accountability measures had already been introduced in the early 1990s.[191]

   Literacy Hour was introduced in all primary schools beginning in 1998 as part of the National Literacy Strategy.[192] The National Literacy Project and the National Numeracy Project were not on a national scale at that time. But in 1998, authorities published a national strategy, one that they felt would be technically and politically viable, delineating its plan to improve learning and quality for all.[193]

Education reform program

6. Built upon the early success of the National Literacy Project, the National Literacy Strategy's goal was to improve upon the low levels of reading and writing by offering teachers a more structured approach to the English curriculum.[194] The approach was based upon the national curriculum that was introduced in schools in 1998. The curriculum described what needed to be taught in schools, the standards of different stages and the sequence, and the recommended time to allow for core subjects.

7. The National Literacy Strategy was more prescriptive than previous reform efforts. It sought to improve the school management of literacy by setting targets, implementing a structured program, and linking these efforts to regularly monitoring and evaluation.[195] The key to the reform was redesigning how teachers taught in the classroom.

   Some of the specific activities introduced by the new policy included the following: making school inspections reflect the new curriculum's aims, introducing league tables of school performance, setting targets for national and individual schools, linking performance to teacher pay, modifying the teacher development program, employing literacy consultants to work with

[189] Machin and McNally (2008).
[190] Stannard and Huxford (2007).
[191] Machin and McNally (2008).
[192] Machin and McNally (2008).
[193] The numeracy strategy was introduced in 1999.
[194] Machin and McNally (2008).
[195] Machin and McNally (2008).

local school administrators and teachers to implement the new strategy, and providing adequate funding to local governments to support them implement the program.[196] While the strategy has evolved over time, the program has improved primary school learning outcomes and narrowed gender gaps significantly.[197]

8. At the school level, the National Learning Program focused on offering teachers specific guidance. This included "term-by-term objectives for teachers, a common approach to teaching through the Literacy Hour, a team of trained and expert consultants dedicated to training and support for all teachers within the common framework plus symmetrical support for head teachers and language coordinators focused on leading teaching and learning, to help them break out of the closed cycle of short-term management and day-to-day survival."[198]

9. One of the most important components of the National Learning Program was Literacy Hour. Teachers were provided with a set of objectives for 5-to-11-year-olds on how to structure time and class management.[199] As Machin and McNally[200] report, "The daily literacy hour is divided between 10–15 minutes of whole-class reading or writing; 10–15 minutes of whole-class session on word work (phonics, spelling and vocabulary) and sentence work (grammar and punctuation); 25–30 minutes of directed group activities (on aspects of writing or reading); and a plenary session at the end for pupils to revisit the objectives of the lesson, reflect on what they have learned and consider what they need to do next."

10. According to the architects of the policy, while the New Labour party was explicitly focused on literacy, their ultimate goal was more ambitious. Through the program, the government sought to transform teaching by applying the principles of the program to broader teaching practices.[201]

Motivations/drivers of education reform

11. The key motivation for the reform concerned the low levels of learning. In 1997, the poor quality of education outcomes was an important issue in national elections. In 1996, the government's Ofsted published a report detailing the low learning standards in schools within three of London's Local Education Authorities.[202] The report traced many of the problems to issues with school leadership. The release of the report itself was a political maneuver. At the time, the Conservatives were in power, but the report detailed learning outcomes of schools that were predominately under the jurisdiction of the New Labour party. When New Labour

[196] Several different approaches have been undertaken to assess the impact of the aspects of the literacy and numeracy program. They are summarized in Machin and McNally (2008); McNally (2015); Stannard and Huxford (2007); Tymms (2004).
[197] Stannard and Huxford (2007); Cassen, McNally, and Vignoles (2015).
[198] Stannard and Huxford (2007, p. 8).
[199] Machin and McNally (2008).
[200] Machin and McNally (2008, p. 1444).
[201] Stannard and Huxford (2007).
[202] Ofsted (1996).

won the election, they had a stated commitment and mandate to assume responsibility for the education system and devise a strategy to improving learning.

12. The government system highlighted poor outcomes that had become a national issue before and during elections. This increased public pressure for change. As the reforms were implemented, better information on school outcomes sustained pressure on schools from parents. The gains came about from a combination of a good political strategy to build support for reforms and a good technical solution.

Reform process

Implementation

13. There was a detailed plan to implement the National Learning Strategy. The reforms addressed many different aspects of learning in a coherent way. For example, the curriculum and teaching strategies to improve literacy and numeracy were developed. Teachers were provided the support to gain the skills they needed to improve their teaching and were motivated to do that through stronger links between performance and pay as well as stronger accountability through information. The reforms were implemented in a sweeping package. There was some experimentation, but they were introduced all together at the same time despite concerns about the capacity to implement.[203]

14. The policy gave some uniformity to how teaching literacy was done across all schools in England. Through this process, Literacy Hour gave teachers a shared pedagogical language for how to talk about and improve education quality.[204] The messaging was also clearer for the public. Since it was a national strategy, parents could know what to expect from schools and their children's teachers.[205]

Information and knowledge

15. The government used data to drive a sense of urgency to improve learning. Ofsted had previously established monitoring and evaluation measures. It was then up to whichever party was in power, be it Conservatives or New Labour, to find ways to improve children's learning.

16. Information on learning was an important aspect of the reform. The Nation Learning Strategy was strongly informed by the documented success of its predecessor, the National Learning Project, which was shown to be effective through case control studies.[206] Evaluation was also used to refine the program. In addition, information came from many different sources, reducing any possibility of bias.[207]

Coalitions and incentives

[203] Stannard and Huxford (2007).
[204] Stannard and Huxford (2007).
[205] Stannard and Huxford (2007).
[206] Machin and McNally (2004).
[207] Machin and McNally (2008).

17. As the architects of the National Learning Strategy, New Labour needed the strategy to work if it wanted to remain in power. Early grade learning outcomes were an important part of the election. New Labour staked some of their political legitimacy and viability on the belief that they could do better at improving learning compared to other political parties.

18. The Local Education Authorities were another key coalition. The schools and teachers under their jurisdiction were the ones responsible for implementing the new policy. New Labour worked closely with Local Education Authorities. Part of the policy included the provision of training, funding, and local consulting so that local educators would not be overburdened.

19. Key to the education sector was the establishment of a task force to develop a comprehensive strategy on early grade learning**.** The task force was led by Michael Barber, whose book, *The Learning Game*,[208] helped inspire the education policies that New Labour adopted. After wide-ranging consultations with other stakeholders, the government published its national strategy.

Innovation and agility

20. The approach and model for reform came about through the National Literacy and Numeracy projects, which included many local education authorities. Though not an official pilot, this allowed different approaches to develop, test, and modify the reforms. The projects in different authorities seemed to coalesce around similar solutions about the best practices for improving learning outcomes.

21. The policy also seemed to be well managed. National policy was developed by learning from the experience of Local Education Authorities while also drawing on existing data and evidence. To implement the policy without placing an undue burden on schools, local consultants were hired. They had the specific responsibility to support implementation at all different levels of the system by working with Local Education Authorities, schools, teachers, and others.

Lessons learned

22. To introduce the policy reform, the political milieu presented a combination of evidence, opportunity, and public support. The New Labour partly came into power on a pledge to improve education quality, and while they already had the political capital to introduce major quality reform, they built on an existing program (National Learning Project), one with demonstrated efficacy.[209]

23. The reform was supported by prior evidence of effectiveness, along with further and multiple evaluations. “Finally, we show the benefits of the Literacy Hour to exceed the costs of the policy by quite a large margin. These findings are of considerable significance when placed into the wider education debate about what works best in schools for improving pupil performance.”

[208] Barber (1996).
[209] Machin and McNally (2008).

24. A key component of the reform, Literacy Hour, was structured and relatively simple to implement across all schools. The program entailed working with Local Education Authorities, but it did not place undue burden on them. The reform enlisted its own consultants and provided funding for training. Learning Hour reconstituted class time that was already intended to be used for teaching English and reading as part of the curriculum.[210]

25. There were no major changes in political leadership during the reform period. Having relative continuity over an extended period of time allowed the education reform to mature. The government could modify the program over time, based upon lessons learned and regular monitoring and evaluation activities.

26. The government was a client to good learning outcomes and public approval rather than other competing factions that were not aligned with improving literacy and numeracy. Low learning outcomes seemed to be the key and motivating driver in the government's commitment to delivering an effective reform. Existing literature does not show that competing coalitions were not aligned with improving learning. For example, Local Education Authorities and teachers' unions did not seem to wield a negative influence when it came to implementing the reforms.

[210] Machin and McNally (2004).

Country case 4: Teacher reforms in Peru, 2007–2012

Introduction

1. When it comes to introducing and implementing teaching reforms, politics matters. Peru's government recognized the role that teachers play in the provision of quality education. In 2007, the government introduced a teachers' reform called Ley de la Carrera Publica Magisterial (CPM) (later called Ley de Reforma Magisterial). The purpose of CPM was to evaluate teachers' competence in pedagogy and subject knowledge. The government ran into opposition by the teachers' union and eventually forced them to make concessions.

   But the government persisted, relying on a public messaging campaign, effective planning, and strategic coalition building. The CPM was one of the factors that probably contributed to improved learning outcomes that Peru has experienced since 2007. The case study highlights the ways in which governments can work to successfully navigate complex political opposition to bring about education quality reforms.[211]

Data/evidence: Trends in learning outcomes and definition of reform improvement

2. Peru has participated in the PISA since 2000. During this time, it has scored below average in reading and math compared to its OECD counterparts. However, the trend in learning outcomes has been largely positive (Figure A5). Between 2000 and 2012, PISA scores in reading improved by one full school year (327 to 384 points), while scores in math between 2009 and 2015 increased by about half a school year (365 to 387 points).[212] On national assessments implemented between 2007 and 2012, the percentage of second-grade students performing at a satisfactory level[213] in reading and math doubled from 15.9% to 30% and 7.2% to 13%, respectively. The percentage of students scoring in the lowest category decreased from 29% to 19% in reading and from 56% to 49% in math.

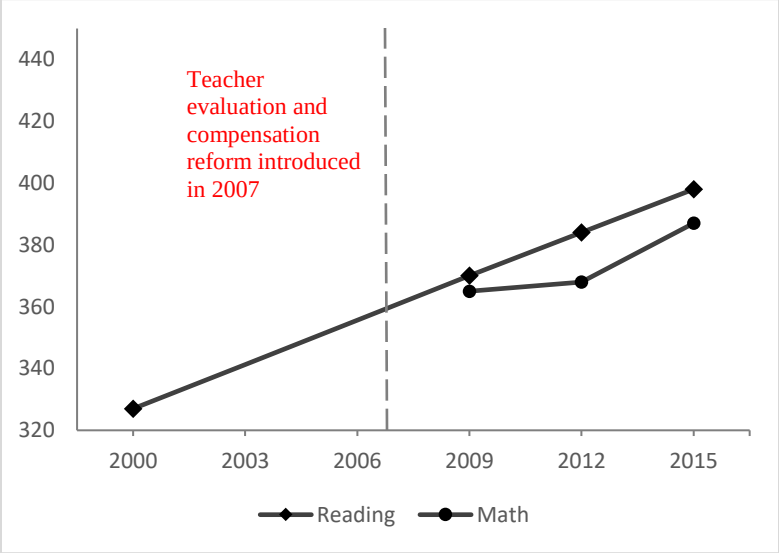
3. Education funding during the reform period surged. In 2003, the national public budget for education was 1.71 million USD. By 2011, it reached 4,600 million USD.[214]

4. While learning outcomes have improved since the CPM was introduced, there are several other ways to explain the PISA scores in recent years. For example, recent research has pointed to a correlation between sixth-grade math students and sixth-grade teachers' performances on the same exam in Peru.[215] About 84% of math teachers performed below Level 2 and close to 48% of teachers below Level 2 in reading. Another more comprehensive test of teachers' mastery

[211] Bruns and Luque (2014).
[212] A forty-point improvement in PISA scores equals one school years' worth of learning.
[213] National assessment results are reported in three levels: "satisfactory" or Level 2, meaning a level of achievement expected by the grade; "in progress" or Level 1, meaning students can only complete very basic tasks; and "initial," below Level 1.
[214] UNESCO (2013).
[215] Metzler and Woessmann (2012).

of pedagogical practice found just 8,000 out of the 183,000 teachers achieved the Ministry of Education's threshold score.[216]

Figure A5: Mean PISA scores in math and reading over time



Source: OECD (2016). PISA stands for Program for International Student Assessment.

Broader context: Overall political and economic context

5. Peru's efforts to implement teacher reforms have been historically difficult, stymied by economic instability and special interests. In 1984, under President Fernando Belaunde, the government introduced a policy to keep the teaching workforce as it sought to expand the provision of education across the country. The policy, Ley del Profesorado, guaranteed job security and introduced seniority-based promotions and wage increases for all public teachers. But later governments have sought to shift to hire and reward teachers based upon qualifications and merit.

6. In the 1990s, under the Fujimori government, the Ministry of Education supported the proliferation of pre-service teacher training institutes, Institutos Superiores Pedgogicos (ISPs). By 2003, 75% of all teacher education enrollments occurred within ISPs. But a 2004 study by the government found that ISPs were not providing quality education to teachers. They were also producing too many teachers than were needed to replace the retired ones.[217]

7. The most significant barrier the government has faced has been a powerful teachers' union called Sindicato Unitario de Trabajadores en la Educación del Peru (SUTEP). SUTEP includes membership from all teaching levels. The union pressurized the government to increase teacher salaries, and it led teachers strikes and blockades until the government gave into their demands. As a result, between 2001 and 2005, teacher salaries increased by 45%.[218]

[216] Bruns and Luque (2014).
[217] Bruns and Luque (2014).
[218] Bruns and Luque (2014).

Education reform program

8. In the midst of this heightened concern about low quality teachers, and strong opposition by the teacher union toward any teacher evaluation reform, President Alan García introduced CPM in 2007. Under this reform, new teachers were required to participate in competency exams. For in-service teachers, the exams were to be voluntary but incentivized through wage increases. Under the CPM, in-service teachers were also required to participate in regular evaluations.

9. In 2008, the government aimed to improve teacher quality by setting higher standards for entrance into teacher training colleges or ISPs. Entry into teaching was now a two-step examination process. The "national phase" of the examination tested candidates' subject knowledge and general skillset. The "regional phase" was administered by local authorities through written tests and interviews. Salaries were set according to performance.

10. The teacher evaluation assigned teachers to one of five CPM levels. Levels were connected with salary increases. Level 1 teachers were given 50% higher salaries than the prevailing wage rate. The salary of Level 5 teachers was twice that of Level 1 teachers. After three years of teaching, Level 2 teachers were eligible for a 15% salary increase.

11. In 2012, the ministry decentralized the admission process by giving control of the selection of students back to ISPs but with an annual enrollment cap set by the ministry.[219]

12. The career system for teachers made implementing the reform difficult for the Ministry of Education. It had to manage the evaluation, promotion, and sanctions of 25% of new hires and in-service teachers. The other 75% opted to follow the previous career path. They received no in-service evaluation but had job security and reliable wage increases based upon seniority. By 2011, just 10% of teachers had voluntarily opted into the in-service evaluation. In other words, the incentives of the CPM were not strong enough to persuade teachers.

13. The CPM was not working as well as it was intended. It was not well received by new and old teachers, and from an administrative perspective, it also proved to be too complicated for the ministry. In 2012, the government introduced a revised version of the CPM, called LRM, which included several components. First, LRM required teachers to give the competency test. Second, teachers could be dismissed after two consecutive inadequate performance reviews or if they had a criminal conviction. Third, the law extended career paths from five to eight levels, and levels were accompanied by salary increases. Fourth, the LRM assessed teacher progression though a holistic approach known as 360 feedback. It included classroom observation by trained professionals as well as feedback from school directors, peers, parents, and community members.[220]

14. One of the factors that reduced the effectiveness of the CPM is that the most important stakeholder was excluded from its development: teachers. For this reason, the government's

[219] Bruns and Luque (2014).
[220] Bruns and Luque (2014).

approach to the CPM was confrontational. The García government was able to push through CPM legislation, but this did not mean teachers were open to receiving its implementation. SUTEP's legacy of violence and disruption meant the government avoided any negotiation with teachers. Instead, the government opted to develop political and communication strategies to garner support of the CPM rather than focus on how to involve teachers in the planning and execution of the reform. Such a situation reduced chances of dialog with the teachers during the planning stage and caused uncertainty due to lack of consensus about the actual reform outcomes.[221]

15. There were several operational and political issues in the CPM reform process that should have been addressed at the time of implementation. The government implemented comprehensive teacher reforms too quickly. It did not offer prior training to improve the weak technical capacity of the Ministry of Education. A third party with limited experience in test design actually designed the competency test for teachers, resulting in major flaws in the testing instrument.[222]

16. There were also some political issues. The national government did not consult regional and municipal governments in the planning stage although local governments bore the economic costs of the program. The federal level transfers to regional government were insufficient to cover the higher salaries under CPM. In addition, the reform implementation was incomplete. Only eight rounds of evaluation were conducted on new teachers, but there were no mandatory performance evaluation of service teachers under CPM. Even the school committees responsible for the second round of teacher selection did not receive the required training.[223]

17. Despite these challenges, teacher evaluation reform stood out because of the effective communication strategy adopted by President García's government. It held public rallies the message of pro-growth agenda and consequent importance of quality education promoted that resonated well with the business community and parents associations. By 2007, a national poll showed that 43% of citizens believed that less qualified teachers were the main reason for poor learning outcomes.[224]

Motivations/drivers of education reform

18. The motivation for each political party to promote or revise the CPM reform was different. The reform aimed to address low subject knowledge in teachers, poor pedagogical skills, inadequate income levels, and a regulatory environment that did not promote professionalism. President García's government in 2007 also recognized the need to leverage the political capital that he gained after his election victory that year. He was motivated to introduce a highly visible, sweeping reform that also aligned with a pro-growth agenda.

    For Humala's government, the drive was mostly to diminish the power of SUTEP. He introduced teacher evaluation strategies (360 feedback mechanism) but also instituted a

[221] Bruns and Luque (2014).
[222] Rivero (2010).
[223] Bruns and Luque (2014).
[224] Cuenca (2017).

staggered pay scale and performance level for teachers. In such a system the union members could not be forceful and cohesive in demanding a similar increase in wages, as all the teachers ended up in different bands of pay-and-performance levels according to LRM. [225]

Reform process

19. Since the inception of reform, the government of Peru had fended off opposition from major power players in the country. SUTEP consistently resorted to political tactics of strikes, public demonstrations, and legal action against the García government to block the teacher reform, as it threatened job security. In return, the government continued to confront the SUTEP's resistance with stronger policy action, responding to union threats by cancelling salaries of all teachers associated with SUTEP. Later, the union challenged the constitutional basis of CPM policy in the Supreme Court, leading to the delay of the implementation of the first round of teacher evaluations. The actions taken by SUTEP persuaded the government to make concession by making CPM mandatory for newly hired teachers but voluntary for in-service teachers.

Information and knowledge

20. Effective communication strategy and showcasing of relevant information helped the government mobilize support from various stakeholders. For example, the SUTEP launched a strike to oppose the merit-based teacher pay system. Members of academia and civil society were also unsupportive of the reform.[226] To get more public support to improve teaching, the government responded by publicizing the poor results of sixth-grade teachers in math and reading. President García also tried to gain public support by holding public rallies in opposition to SUTEP. A poll showed that 74% of people were convinced that CPM would be beneficial for the general public and teachers.

21. Despite the having political backing of the union in the 2011 elections, the new government, under President Ollanta Humala, continued with a more modified CPM law to reform teachers' development and evaluation practices. The reform was supported due to Education Minister Jaime Saavedra's convincing communication style when promoting the reforms' effectiveness in the media.[227]

Coalitions and incentives

22. President García's government used its political capital to confront the opposing parties and gather support from a wide range of stakeholders to promote its reform agenda. The government used a strong media strategy to get the support from the Business Association for Education (BAE). Under President García, the BAE funded a publicity campaign to promote the importance of quality education for economic growth.

    The National Parents Association was also an important stakeholder. Once it supported the CPM reform, it advocated for the ousting of teachers who were participating in SUTEP-led strikes and protests. Other stakeholders included NGOs, who advocated for education policies

[225] Bruns and Luque (2014).
[226] Cuenca (2017).
[227] Cuenca (2017); Bruns and Luque (2014); Bruns and Schneider (2016).

during this time,[228] and technocrats, who helped improve later iterations of the reform. Under Humala's government, stakeholders included economic and planning ministries, from whom the Minister of Education solicited expertise to improve the reform.

Innovation and agility

23. The CPM aimed to improve teacher effectiveness through offering a simple, mandatory, and streamlined approach. Since its inception, SUTEP has been a barrier to improving education quality. But President García decided that, rather than cede to its demands like his predecessors had done, he would confront the powerful union. He used an effective communications strategy to gain support from key stakeholders (except teachers), and this approach diminished the power of SUTEP.

    In 2012, under the Humala government, the revised CPM, called LRM, made it mandatory for all teachers to go through the evaluation process on a regular basis, offering high salaries to motivate teachers. Streamlining the CPM through the LRM reform made implementing and managing the reforms much easier for government. Through its 360 feedback program, teachers benefited from written feedback, along with input from school directors, peers, and parents.[229]

Lessons learned

24. Amend teacher reforms based on past performance: Instead of completely scrapping the original CPM reform, the new government studied the initial impact and made targeted amendments to the existing reform program for better impact. When it was introduced, the merit-based teacher development reform required in-service teachers be evaluated on a voluntary basis. Even though the evaluations came with financial incentives, teacher uptake was low. The new iteration of the CPM reform, LRM, mandated all in-service teachers to undergo evaluations regularly, and pay scales were based on exam performances.

25. Use popular support and momentum to confront powerful teachers' unions: President García used his political capital to introduce sweeping teacher reforms early on in his tenure. It was an aggressive move but one that allowed him to leverage his political capital and strong popular support to directly confront SUTEP. While the union delayed the reform process, the government challenged the union by building a broader coalition. They used evidence of poor teacher performance in sixth-grade examinations and low learning outcomes to galvanize public support for introducing new reforms in the country.

26. Alliances and public campaigns can build momentum for reform: The government gained support from key stakeholders, such as the business community and parent associations, in part due to the compelling rhetoric and leadership. Peru's leaders used public rallies to build support and to make the case that the reforms were needed if the country were to develop. The financial and political support provided by the alliances helped the government persevere through contracted opposition and even sustain reform efforts in subsequent regimes.

[228] Bruns and Luque (2014); Bruns and Schneider (2016).
[229] Bruns and Luque (2014); Bruns and Schneider (2016).

Country case 5: Big Results Now Education (BRNEd) program in Tanzania, 2012–2016

Introduction

1. Since achieving near-universal access, Tanzania has endeavored to improve the quality of its basic education system. Between 2012 and 2016, the government of Tanzania set out to improve education quality in a way that was consistent with its national economic development scheme, called Development Vision 2025. Based upon a similar model developed in Malaysia, Tanzania introduced BRNEd in 2013, a multi-sectoral plan to transfer funds to high-priority schools across the country while minimizing the potential for patronage and rent extraction. By the end of the reform period, the government observed gains in student performance, teacher training, and financial management.[230] While the improvements cannot be definitively linked to BRNEd, this case explores the lessons that can be drawn from the reform process in terms of development and implementation.

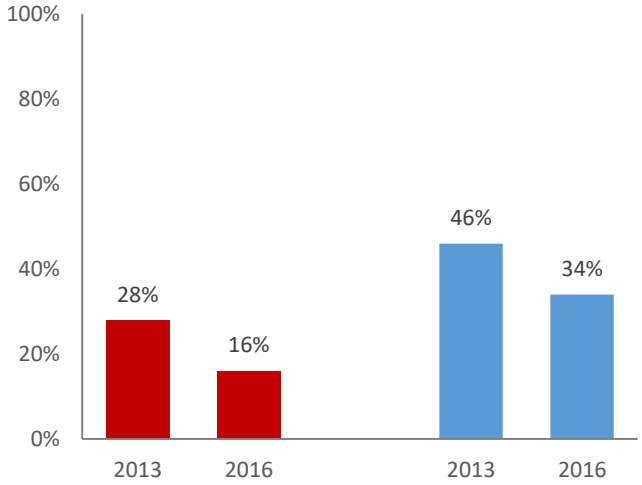
Data/evidence: Trends in learning outcomes and definition of reform episode/period

2. By 2012, access to primary school was nearly universal,[231] but indicators concerning quality lagged. Many students in their final year of primary school failed to acquire the basic literacy and numeracy skills expected of a student in the second grade.[232] About 47% could not read English stories; 26% could not read stories in the local language, Kiswahili; and 11% of the students could not perform basic multiplication.[233] Between 2011 and 2012, primary school pass rates in Tanzania declined from 58% to 30%.[234] And between 2012 and 2014, literacy and numeracy rates were also low.[235]

3. However, by the end of the reform period, indicators of education quality began to show signs of improvement. According to a 2016 mid-term review by the World Bank, Tanzania improved its EGRA and EGMA scores compared to 2013. The share of students who achieved the "progressing reader" level increased by 40%, and the share of those approaching the math benchmark increased by 18%.[236]

[230] NKRA (2014).
[231] UNESCO (2012).
[232] Uwezo (2012); World Bank (2014).
[233] Sabarwal, Joshi, and Blackmon (2016).
[234] World Bank (2014).
[235] Uwezo (2014); Uwezo is a household-based assessment of literacy and numeracy skills for 6-to-16-year-olds. The results are shared by community members and policy makers to prompt policy action connected to the "three Rs" (3Rs): reading, writing, and math.
[236] World Bank (2016).

Figure A6: Share of low performers in early grades, 2013–2015



Source: RTI International (2016).

Broader context: Overall political and economic context

4. In the late 1960s, President Julius Nyerere applied a socialist approach to developing reforms in newly independent Tanzania. His administration centralized and uniformed policies, and they aligned schools' organization and curriculum to provide universal education that prepared students to work to develop their communities. Led by the ruling party, the Tanganyika African National Union (TANU), the government saw education as central to its state-building goals.

   Efforts to universalize primary school began in 1974. By 1981, almost 90% of primary school students were enrolled in schools.[237] In early 2001, the government made primary education free and mandatory. By 2011, primary enrollment reached 94%.[238] But the education system continued to suffer from poor education quality, low learning standards, and an undertrained teaching workforce.[239]

5. TANU has been the ruling party since independence in 1961. While the country's approach to governance has been historically centralized, recent efforts have been made toward devolution. In 1998, the Local Government Reform Program was introduced. This program aimed to empower local levels of government by giving them more autonomy, yet the national government controlled the distribution and allocation of funds, based on strict guidelines and regulations concerning resource allocation. About 90% of the budget was allocated by the government to local authorities, with the remaining 10% collected by local authorities through taxes and levies.[240]

[237] Sumra and Katabaro (2014); James (2014).
[238] NKRA (2014).
[239] Sumra and Katabaro (2014); NKRA (2014).
[240] Mollel and Tollenaar (2013).

6. Between 2008 and 2012, Tanzania's education expenditure as a percentage of GDP was about 5%. In 2013/2014, the government allocated 70% of the education budget spending to district-level governments, and wages and personal emoluments comprised 46% of the education budget. But there was a problem: the government could not equally disburse resources across regions on a per-pupil basis. Inequality in education spending across regions reflected the unequal distribution of teachers across regions and districts. Inefficiencies in education spending also were due to high repetition, high dropout rates, and low pass rates on national examinations.[241]

Education reform program

7. The government introduced BRNEd in 2013, modeled after Malaysia's Big Fat Results program. Tanzania's government has set a goal of becoming a middle-income country by 2025. It aims to achieve its development goals by identifying and monitoring progress of National Key Results Areas (NKRAs) and by devising programs through lab sessions.[242]

8. When BRNEd was developed in 2012, the government organized six week-long lab consultations with 34 members of governmental and nongovernmental organizations. They represented teachers' unions, academia, international aid organizations, government ministries, and local level government officers. Over the six weeks, they deliberated to create the BRNEd reform strategy.[243] Together, they settled on the nine key areas, each of which were intended to rapidly improve education. Given the low quality, the government implemented BRNEd with a sense of urgency. It followed an aggressive timeline, beginning in early 2013 through the end of 2014.[244]

9. BRNEd decided upon six NKRAs, one of which focused on the education sector. To improve education, the government outlined nine key high-impact initiatives that were classified under four strategic levers: (1) transparency to help better assess school and student performance, (2) incentives to encourage better performance, (3) support for teachers and students, and (4) teacher conditions to improve teaching conditions.[245] Table A3 outlines each of the high-impact initiatives.

Table A3. BRNEd nine key high-impact initiatives for the education sector[246]

Transparency	
1. Official school ranking	Primary and secondary schools publicly ranked on performance and improvement in examination pass rates.
2. National three R (3R) assessments	A sample of second-grade students in schools from 140 councils was tested on 3Rs through existing assessments (EGRA, EGMA, and Uwezo) to establish a

[241] Sumra and Katabaro (2014).
[242] NKRA (2014); Puttick, Baeck, and Colligan (2014); RISE (n.d.).
[243] RISE (n.d.).
[244] NKRA (2014); RISE (n.d.); Sabarwal, Joshi, and Blackmon (2016).
[245] NKRA (2014).
[246] NKRA (2014); RISE (n.d.).

	national baseline for 3R competencies in the country.[247] This assessment program is to expand to all 140 councils across the mainland.
Incentives	
3. School incentive scheme	Schools were annually awarded a certificate of excellence in public ceremonies, based on the students' performance on national examinations. Schools had to show an improvement of 10% or more from the previous year to be eligible for a prize. The best performing schools received awards to improve their school.
Support	
4. School improvement toolkit	Head teachers were offered training on school management.
5. 3R teacher training	Teachers were trained to teach the essentials of 3R and identify students who had fallen behind in basic competency skills.
6. Student Teacher Enrichment Program (STEP)	Primary and secondary teachers were trained to identify and help low-performing students.
7. Basic facilities construction	There was a fast-track construction of 1,200 additional school facilities through streamlined procurement processes, better oversight, and improved coordination between government entities.
8. Capitation grants (CGs)	Reasonably efficient financial support provided through disbursement of earmarked education budget to local governments for buying school resources on a per-student basis. CGs ensured revenue maximization and better monitoring of equal per-student financing to keep in check any leakages of district-level funds.
Teacher conditions	
9. Teacher motivation	Teacher conditions were improved through monetary and nonmonetary rewards. Any outstanding salary and nonsalary claims of teachers were also cleared.

Notes: BRNEd stands for Big Results Now Education, EGRA for Early Grade Reading Assessment, and EGMA for Early Grade Mathematics Assessment.

10. Implementing BRNEd was a challenge for the following reasons. First, the timeline of the program was short, which provided little margin for error in terms of implementation. Second, the reform was bureaucratically complex. It entailed the contributions of several government ministries (Ministry of Education and Vocational Training, prime minister's office, and Ministry of Finance), two new BRNEd implementation units (President's Delivery Bureau and Ministerial Development Unit), and local government authorities. Third, the level of technical capacity among government agencies and local government bodies varied in terms of its ability to successfully implement BRNEd. Fourth, this lack of capacity meant that the government often hired retired government officials to serve as consultants. They were typically paid more than current government payroll staff, and the disparity threatened morale.

    Fifth, the government lacked effective oversight of local government authorities when it came to distributing CGs. Issues of inefficiency and regional disparities were common. As a result, the government funded only 56% of the committed CGs during fiscal year 2013–2014, but some districts ended up receiving 11 times more than other districts.[248] To avoid problems associated with the capitation grant, the government aimed to transfer CG funds directly to

[247] NKRA (2014).
[248] World Bank (2014).

school accounts. But by 2013, this capacity was not in place yet.[249] Budgetary constraints also limited the awards to be conferred to qualified schools under the school incentive scheme and caused delays in teacher training. Finally, the use of examination results to rank schools may have encouraged teachers to "teach for the test" rather than focus on the conceptual development of skills and competencies in reading and math.[250]

11. To date, no rigorous evaluation of BRNEd on learning outcomes has been implemented. But on a process level, the positive trends in some of the key reform areas serve to highlight how it may have impacted learning. The school incentive scheme resulted in close to 3,000 "best" and "improved" school designations in 2014. This process led to productive discussions among parents and students about how quality could be improved at the school level, which enhanced community-level participation in local decision-making concerning education. The school rankings scheme was also an effective way to hold teachers, school administrators, and district-level governments accountable to improve education quality.

12. The program has also supported teachers. Approximately 2,900 school heads received school management training, 4,000 teachers received STEP training, and 1,300 schools have started STEP classes for low-performing students. The Prime Minister's Office-Regional Administration and Local Government (PMO-RALG) managed to clear 86% of outstanding teacher claims. This financial reform positively impacted the persistently low teachers' morale and improved teachers' confidence in government reform efforts. Also, the results of 3R assessments of 2,000 Standard II students in 200 schools in 2013 were made public; therefore, teachers could better understand the basic skill levels of their students and the type of teaching methods to apply.[251]

Motivations/drivers of education reform

13. The key motivation behind the BRNEd initiative was to address the issues of low pass rates in the country's two national examinations, the Primary School Leaving Examination (PSLE) and Certificate of Secondary Education Examination (CSEE), which occur at the end of primary and secondary school. Tanzania's government identified critical improvements to increase examination pass rates. Under President Kikwete's leadership in 2012, the government approached the problem with a multi-faceted program aimed to quickly upgrade the skills of students, teachers, and administrators; make managing educational finance more efficient; and set benchmarks to measure learning achievement through strong accountability mechanisms.

14. The decision to focus on education quality was also political. On one hand, the government used BRNEd to respond to a public outcry about poor quality education. On the other hand, the government was simultaneously positioning itself for the upcoming elections in 2015. The Chama Cha Mapinduzi (CCM) political party wanted electoral victory in 2015, and a high-profile education program was one of the ways to win support across the country. The initiative

[249] World Bank (2014).
[250] Sabarwal, Joshi, and Blackmon (2016).
[251] NKRA (2014); World Bank (2016).

would show concrete evidence of the positive impact of Kikwete's leadership. Regional and district commissioners responsible for overseeing the BRNEd programs' implementation process were mostly members of the ruling political party. Thus, their interests and incentives aligned with the successful implementation of BRNEd.

15. The prime minister's office, PMO-RALG, oversaw the activities of the local government administration apparatus (LGA). While the central government devolved power to local authorities, they still based most of the local education plans on national guidelines. Therefore, the role of LGAs in implementing most of the district-level education programs aligned with district and regional government guidelines that aimed to gain more political support from the public.[252]

16. Another motivating factor for the government was to develop the reform in such a way that could attract more foreign aid. It aimed to do so by ensuring that the scope of BRNEd would cover the whole country. The government also aimed to ensure the funds were used well, and they transferred funds directly to schools as a way to avoid leakage.[253]

17. At first, teachers' unions did not support the reform program. They viewed education reforms as another way of giving teachers more responsibilities without enough compensation. To receive the buy-in of unions, the government aimed to clear any salary and nonsalary dues they were owed. In addition, the financial incentives awarded based on teacher performance was another way to gain teacher support of BRNEd. By extension, teacher buy-in also meant it could help the ruling party in the next election. Therefore, more than using reform to increase learning outcomes, the government was motivated to win political support.[254]

Reform process

18. Given the consensus among stakeholders, accountability measures for schools and students, training programs, and salary disbursement system for teachers, BRNEd led to some modest success. For instance, PMO-RALG cleared 86%; 76% of outstanding teacher claims were disbursed in 2014; the percentage of nonreaders dropped from 27% in 2013 to 16% in 2016; and the percentage of nonperformers in EGMA dropped from 46% to 34% between the time BRNEd began and 2016.[255]

Information and knowledge

19. To develop and implement BRNEd, the government extensively relied on data. For example, the declining pass rates in PSLE and CSEE between 2006–2012 were a major cause for concern for the public and government, especially with national elections due in 2015. The scores in international assessments, such as EGRA and EGMA, also highlighted that low education quality was a systemic issue. When the government was developing BRNEd, they did so on the basis of examination data. To monitor any progress, the government set benchmarks for

[252] Mollel and Tollenaar (2013)
[253] RISE (n.d.).
[254] Uwezo (2014).
[255] Sabarwal, Joshi, and Blackmon (2016); NKRA (2016).

national pass rates, 3R assessments, and key performance indicators (KPI). For instance, public schools were ranked and awarded annually according to pass rates and score improvements.

The National Examinations Council of Tanzania (NECTA) is the government body responsible for conducting national examinations and for ranking government schools based on performance. The NECTA has even published an online interactive district-level map, showing performances in PSLE and CSEE examinations in 2013 and 2014 to increase the accountability of school administrators and teaching staff. Government officials were aware of the possible moral hazard due to rankings based on examination results. Therefore, to avoid schools from excluding weak students from the examinations, the government introduced the STEP program (see Table A3) to train teachers to identify weaker students early on and offer them remedial classes. Even after implementing BRNEd, the government continued the results-oriented approach through data analysis. They did this by soliciting support from international aid organizations to develop the Education Management Information Systems (EMIS).[256]

Coalitions and incentives

20. The formation of the BRNEd reform was a team effort by 34 members belonging to 31 diverse groups of organizations. They modeled their discussions on the same approach used in Malaysia. As described above, a wide range of governmental and nongovernmental actors were consulted. A teacher union was one group that had been reluctant in sustaining the reform efforts, mainly because of outstanding teacher claims even after a few years into the reform. Teachers viewed training and skill development opportunities in the reform process as motivating but also a way of piling more responsibilities without any significant monetary benefit. In response to teachers' discontent over unfulfilled claims, the government argued that most of the claims were illegitimate because they could not be verified.[257] However, government efforts to motivate teachers through monetary and nonmonetary rewards were enough to win over the teachers' union to support BNRED.

Innovation and agility

21. Forming the reform process involved what was called the "lab consultation" method. This method allows all stakeholders to share a physical space to deliberate and receive feedback on how to resolve a problem at hand—in this case, education quality. The lab produced a lab report that outlines their agreed-upon plan. The government of Tanzania created two independent units: the President's Delivery Bureau (PDB) and the Ministerial Delivery Units to help expedite the reform implementation. In 2013, the president himself mandated that the PDB directly oversee implementing NKRA according to the established KPIs. The officials from the president's office were also part of the 32-member coalition that devised the BRNEd initiative in 2012.

    The president, perhaps aiming to win political support for the upcoming election through encouraging BNRED results, held assessment meetings every six months to track the implementation progress. In addition, he appointed a third party to assess the performance of ministries and district offices with respect to the established KPIs. This setup ensured that the reform process would be efficient, and it continued accountability of the stakeholders. Apart

[256] World Bank (2014).
[257] RISE (n.d.).

from soliciting funds from international donor organizations, the government planned to adopt an unconventional funding approach to source funds for the education sector. To do this, they created the Education Investment Levy, which taxed calls made with mobile phones. The government also decided to ringfence the capitation budget so that there is timely and regular flow of CG to schools directly instead of through LGAs.[258] In the future, the government plans to tap into private sector debt markets through issuing Eurobonds.

Lessons learned

22. Commitment from the higher office keeps the reform momentum strong. Different reform programs implemented under BNREd during 2013–2015 continued beyond the official completion of its implementation. President Kikwete continued with the governance system that promoted single-party rule through a strong central government. Partly encouraged by political motivations for the 2015 elections, the president showed keen interest in ensuring effective delivery of reforms. The president's office's direct involvement in the lab consultations during the formation of the reform strategy, and the later establishment of the PDB to oversee BRNEd's implementation process, signals the government's strong commitment for education reform. Regular presidential performance dialog to check the progress of different sectors and line ministries ensures that the reform process continues smoothly and any roadblocks are cleared early on.

23. A unique and inclusive consultation process can produce a comprehensive reform strategy. The lab used a syndication process whereby proposed ideas and solutions were shared with other consultants in the lab for feedback. After the general consensus was reached on a possible reform, the results were published in a lab report, detailing responsibilities of the reform implementation plan. The diverse and relevant stakeholders involved managed to conceive the BRNEd policy that aligned with broader education goals of the country, as outlined in Vision 2025. In addition, the processes for executing and monitoring the battery of programs under the reform were clearly outlined. The inclusive nature of the initial consultation process could lead to the achievement of the end Vision while holding the relevant stakeholders accountable.

24. Accountability can improve the quality of education by making reform results public knowledge. Sharing official school rankings, based on student performance in primary and secondary examinations, pushed the school management to work harder and offer a quality learning experience to students. Public knowledge also triggered constructive discussion among the community members to question the administration of nonperforming schools and put a spotlight on teachers who have shown dedication to their profession.

    Uwezo was first conducted in the country in 2010. The results of the latest assessment, in 2012, were shared with the community members, parents, teachers, school administrators, and policy makers to influence education reform agenda. The dismal literacy and numeracy results of 2012 Uwezo, including discouraging results from other assessments, prompted the Tanzanian government to begin BRNEd.[259]

[258] RISE (n.d.); Government of the United Republic of Tanzania (2013); NKRA (2014).
[259] Uwezo (2014).

Data can help policy makers make decisions for the education sector. For example, falling pass rates in examinations alarmed the government about the deteriorating state of the education system in the country. Later governments continued to use quantifiable metrics, such as EGRA and EGMA, to find out the efficacy of BRNEd reform activities in schools and for students and teachers. The nine initiatives under the education NKRA were assessed based on KPIs for each reform action; those numbers also revealed the implementation capacity of the officials responsible. Even after the BRNEd program officially concluded, the government realized the need to remove inefficiencies in the system related to teacher deployment, grants disbursement, and student testing. They sought support from the World Bank to develop the EMIS system so they could better assess the performance of complex policy programs.

Country case 6: Aligning curriculum reforms in South Africa

Introduction

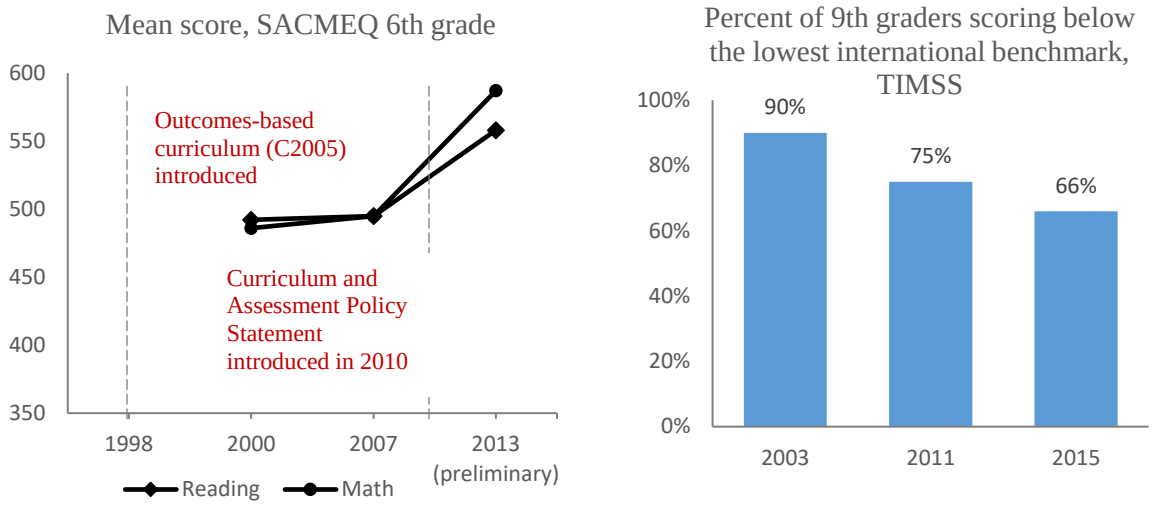
1. This case study investigates South Africa's efforts to improve education quality through the curriculum reforms that took place between 1994 and 2015. South Africa introduced its first national curriculum following apartheid in 1994. In the years after it was launched, learning outcomes remained low. But through a process of modification and realignment over the next decade and a half, marked improvements were made on national assessments. We investigate why this was the case by exploring the political and economic drivers of curriculum development and reforms.

Data/evidence: Trends in learning outcomes and definition of reform episode/period

2. In the years immediately following the end of apartheid, primary and secondary school learning outcomes were low. The first iteration of the country's national curriculum occurred in the mid-1990s. But by 1999, progress was not yet seen on international learning metrics. For example, among the 12 African countries that participated in the Monitoring Learning Achievement[260] project, South African fourth graders scored the lowest in numeracy, the fifth lowest in literacy, and the third lowest in life skills. TIMSS scores[261] also showed stagnation for the lowest achievers between 1996 and 1999.[262]

Figure A7: Performance in cross-national assessments over time

[260] Chinapah (2003).
[261] TIMSS (2015).
[262] Eighth-grade students appeared in TIMSS in 1996 and 1999 from South Africa.



Sources: SACMEQ results are from the South Africa Portfolio Committee on Basic Education (2016); TIMSS results are from Reddy et al. (2016). SACMEQ stands for Southern and Eastern Africa Consortium for Monitoring Educational Quality, TIMSS for Trends in International Mathematics and Science, and C2005 for "Curriculum 2005."

3. However, learning outcomes have improved more recently. While performance of sixth graders in the SACMEQ tests was relatively stagnant between 2000 and 2007, preliminary results for 2013 suggest an improvement (Figure A7). Math performance at the secondary level also increased compared to the early 2000s. Between 2002 and 2011, students scoring in the highest achievement levels on the TIMSS increased, and students scoring in the lowest achievement levels decreased for ninth graders. Also during that time period, students in the lowest level of math achievement decreased by 15 percentage points; the lowest performers in science fell by 12 percentage points; and in 2011, the top performers in math rose by 0.4 percentage points.

Broader context: Overall political and economic context

4. The end of apartheid ushered in a period of immense social, political, and economic transition for the country. In 1995, Nelson Mandela's ANC party won on a platform that championed racial equality, including the introduction of an integrated and equitable national education system. The government first concentrated its effort on improving access to education. By the time of independence, primary net enrollment rate had reached nearly 95%.[263]

5. The ANC also supported a link between economic growth and developing a fair and just society, one that could be used to overcome its legacy of inequality. As such, the government wanted to strengthen education and improve the quality of human capital to develop a competent and qualified workforce as well as to introduce a national curriculum.

6. The changes that the government made during this time also occurred during a shifting global discourse about the state's role in the provision of basic education. There was a growing realization that the structural adjustment programs of the 1980s had failed to bring sustainable

[263] Chisholm and Leyendecker (2008); UNESCO (2012).

change for poor countries like South Africa. In its place, democracy was now seen as a necessary ingredient for economic development, fueled by an educated and engaged citizenry.[264] Before the ANC, South Africa had already become a signatory to the World Declaration on Education for All. But the new government's pledge was more than a visible commitment to the global community. It also symbolized the ANC's commitment to providing a sustainable and inclusive education in a way that would benefit children across the country.[265]

Education reform program

7. The government's effort to reform the curriculum can be characterized as iterative. The country has participated in the TIMSS assessments since 1995. These international metrics were considered to be a reliable litmus test of the state of children's education, as well as a tool that could be used to improve the education system, including the curriculum. Thus, when test scores did not improve, the government altered the curriculum.

8. The most significant reform to the curriculum under the ANC was the introduction of a country-wide curriculum. Before the ANC, districts had their own curriculum. But in 1996, the first post-apartheid Minister of Education, Sibusiso Bengu, announced the National Education Policy Act. The act established a national education system to be used by the whole country.[266] It also decentralized the national system, giving provinces the autonomy to deliver the education and training in ways they needed.

9. Major reforms to the curriculum can be divided into three phases. The first phase was the introduction of the curriculum in 1997, known as "Curriculum 2005" or "C2005." To develop C2005, the government worked to consolidate 19 curricula into a single national curriculum, one absent of racist language. They began implementing C2005 in 1998; it consisted of 8 learning areas and 66 learning outcomes.[267]

   The introduction of C2005 along with the curriculum revisions that followed were predicated on the international "outcomes-based education" approach. One of the major changes of the curriculum was that "subjects" were replaced by "learning areas," and "competencies" were reframed as "outcomes." Teachers were instructed to start tasks as outlined in the new curriculum. The challenge, however, was that while teachers were told what to teach, they were not told how to teach it. The government placed little emphasis on equipping teachers with the pedagogical skills they needed to be effective in the classroom.[268] Similarly, students were being asked to learn in new and innovative ways, but they too did not yet possess an analytical skillset that would enable them to learn through a student-centered curriculum. As a result, despite efforts to improve learning, test scores remained low compared to other African countries.[269]

[264] OECD (2008).
[265] Chisholm and Leyendecker (2008).
[266] OECD (2008).
[267] OECD (2008).
[268] OECD (2008).
[269] Jansen (1998); Chisholm and Leyendecker (2008).

10. The second phase of curriculum reform took place between 2000 and 2002. The government formed a committee to assess the issues in implementing C2005. The challenges the committee identified included limited teacher roles, high student expectations, dependence on limited materials in schools, minimum provincial support, and an unrealistic timeframe to implement C2005.[270] The committee produced a Revised National Curriculum Statement, which led to important changes to C2005. The revision placed more emphasis on content knowledge and basic skills, and it developed a more logical approach toward class progression so that the curriculum identified learning goals for each class. The reforms introduced 8 learning outcomes in lower grades and 29 outcomes for higher grades, and they also introduced technology as a school subject. The reforms also placed more emphasis on teacher training.

    In 2000, the South African Council of Educators was established for the professional development of teachers. Since 2004, about 2,100 provincial officials and 200,000 educators went through an orientation program.[271] But despite these modifications, high scores in international and regional assessments remained low. One explanation was that the teachers were left to define the content of the subjects for themselves and the students' assessment was based on vaguely defined outcomes.

11. The third phase occurred between 2010 and 2014. In 2010, the government introduced a Curriculum and Assessment Policy Statement (CAPS). CAPS was designed to address some of the ambiguity in prior iterations of the curriculum. It provided teachers with details about what topics and subtopics they needed to teach in each class. The goal was to provide more uniformity to the national curriculum rather than leaving it up to the teachers to figure out what and how to teach on their own. CAPS also emphasized delivering skills to help young people transition from school to work. It also included a stronger emphasis on South Africa's history and heritage.

    Students across the country were also given the same national assessment on an annual basis.[272] This added structure and guidance now offered by the curriculum, and it benefited many educators. While it came at the cost of constraining teachers' autonomy of what and how to teach, those from disadvantaged backgrounds benefited because it gave them further guidance on exactly what to teach, even if they lacked experience compared to the better-off teachers.[273]

12. Taken together, we can see that South Africa took incremental and iterative steps to improve student learning. To be sure, the introduction of a new and more structured curriculum in the late 1990s was imperfect, but modifications over the next 15 years, such as better guidance to improve the delivery of the subject areas, probably contributed to improved learning outcomes that occured later.

Motivations/drivers of education reform

[270] OECD (2008); Cross, Mungadi, and Rouhani (2002).
[271] OECD (2008).
[272] Department of Basic Education (2018).
[273] Msibi and Mchunu (2013); Ramatlapana and Makonye (2012).

13. The ANC came into power at a time when South Africa was deeply divided along racial lines and class inequalities. The new government was eager to develop the basic education system in a way that could symbolize its commitment to all South Africans. Their introduction of the C2005 was designed to eliminate racism and promote national unity, introduce a learner-centered approach to teaching and learning, and improve the country's human capital to advance economic development.[274]

14. Some scholars have noted that the new curriculum's introduction was as much political as it was developmental.[275] This introduction symbolized the government's commitment to improving the education system—even if the changes were largely made to immediately improve learning outcomes for children.[276] In short, moving quickly to reform the education system through a new curriculum was an act of political expediency, even though the curriculum lacked content and direction on how to have teachers deliver it effectively.[277]

Reform process

15. The policy experience of the apartheid-era government, the GNU, allowed them to effect policy reforms because unlike the newly elected ANC, the GNU had experience in policy making. Then GNU members in the new coalition government tried to promote the interests of white South Africans by advocating for continuing apartheid-era policies.[278] For example, they helped implement a new, complex curriculum without any prior teacher training and development program, but they did support school management by inexperienced and sometimes uneducated parents through the SGBs. However, this had the effect of reinforcing racial inequalities. Most white children attended white schools whose teachers were already well trained in pedagogy, while most black children attended schools where teachers had not received adequate training.[279]

16. White factions continued to impact education policy in other ways, too. For example, teachers from elite white schools were part of the learning area committees that provided consultations right after independence.[280] The reforms that took place under the new curriculum had disproportionately negative effects on poor provinces that mostly served black South Africans. Black students in poor provinces eventually attended poor performing schools, which resulted in students performing at the lowest tier of achievement on learning metrics.

17. However, the curriculum reforms made since 2000 slowly has improved education quality. For example, based on the students in quintiles[281] 1, 2, and 3 from the poorest "no-fee" schools, comprised mostly of black students, scored the least on TIMSS 2011 compared to the affluent

[274] Cross, Mungadi, and Rouhani (2002).
[275] Cross, Mungadi, and Rouhani (2002).
[276] Cross, Mungadi, and Rouhani (2002).
[277] Cross, Mungadi, and Rouhani (2002).
[278] Cross, Mungadi, and Rouhani (2002).
[279] OECD (2008).
[280] Govender (2015); Chisholm (2005a, 2005b).
[281] The quintile rank for each school is calculated based on the poverty level of the community in which the school is located. Quintiles 1, 2, and 3 were declared no-fee schools in 2010.

quintile 4 and 5 schools. But the poorest quintile students also showed improvement between the 2002 and 2011 TIMSS, and a small percentage of students even entered high and advanced achievement levels.[282]

18. In 1994, teachers' unions were influential in shaping curriculum reform. Teachers' unions were legalized by the government. However, their role advocated for some reforms but not others. For instance, teachers' unions had little influence on the curriculum itself because the government chose to align the curriculum content with the interests of the private sector, such as trade and labor unions, in the hopes of stimulating economic growth. The teachers' unions shifted focus to advocacy on higher salaries and better work conditions rather than learning.[283] They were also influential in limiting the role of parents in advocating for national education policies.

    While unions can influence national policy to some degree, the role of advocacy by parents has been limited to the subnational level through the introduction of SGBs. More recently, however, the national teacher union, the South African Democratic Teachers Union (SADTU), has started to shift its focus toward teacher development. The SADTU established the Curtis Nkondo Professional Development Institute in 2013.

19. During the second revision of C2005, the government formed eight working groups of diverse faculties to develop a plan to revise the curriculum. The government was open to incorporate the innovative ideas of economic and management sciences and technology into the revised curriculum. History and environmental lobbies also provided input, which was accepted. A Christian lobby also took place in public hearings, but their efforts were rejected in favor of a secular curriculum.[284]

20. The government also formed review committees during the reform episode. The committees consisted of university intellectuals, politicians, and other key stakeholders. They worked to identify implementation issues that plagued the first iteration of the curriculum and shifted the focus of reforms toward quality learning in schools.[285] During the early phases of policy formation, the ANC made SADTU part of the Education Study Group, which was one of the groups formed in the early part of the reform movement. Its purpose was to have dialog among stakeholders to devise education policies aimed to achieve equity and quality. [286]

Lessons learned

21. Education reforms were political and developmental and were sometimes at odds with one another. The reforms undertaken by the ANC can be understood as seeking to improve social cohesion and promote the economic development of the country. But the reforms also must be understood by examining the elite's interests and incentives, and the approach to the

[282] SACMEQ III (2017); Reddy et al. (2015)
[283] Govender (2015).
[284] Chisholm (2005a, 2005b).
[285] Chisholm (2005a, 2005b).
[286] Chisholm (2005a, 2005b).

curriculum reform must be understood in political context. Two examples from this case study illustrate this point.

First, the ANC introduced the curriculum without providing adequate training for teachers. In other words, the ANC received credit for their efforts to develop a nationwide curriculum, but it had little chance of improving children's learning without any training. Second, we can look at the decentralization and devolutions of powers through the SGBs. SGBs said they intended to empower local education systems to use resources in a way that reflected their needs.

But as we saw in this case study, SGBs received little training or oversight. While educated, white regions could use SGBs, black and historically underserved parts of South Africa could not. This was not a technical oversight. Rather, the absence of training was engineered by powerful pro-white forces and coalitions that saw the SGBs as a chance to reinforce apartheid-era power divisions. As a result, for example, former white schools whose management had prior training in financial management have been performing well and using the funds in a more constructive manner, compared to former black schools, where SGB members need additional training in running the expenses of schools and work more efficiently.[287]

22. Participating in international assessments provided a shared platform and benchmarks of accountability through which to build upon. One important ingredient in the improvements watched over time was the emphasis in assessment. Tests such as the TIMSS and SACMEQ provided reliable and valid data that the government took seriously as an important indicator for the quality of the education system. Using these test results, they revised and introduced various programs within the curriculum reform.

23. To be sure, the reforms to the education system introduced profound change through the country's first national-level curriculum. But in terms of improving the quality of children's learning, we can look back and see that the curriculum was modified, adapted, and envisioned over time. The government saw that children were not learning as well as they would have hoped. To improve delivery, they called panels of experts and stakeholders to revisit and revise aspects of the curriculum and the training they were providing teachers. They did not scrap the curriculum when it failed to lead to wanted outcomes, but it did improve over time. For example, the changes accompanied after the first phase of the reforms focused more on curriculum development and devolving power in education provision.

    However, what was equally important was the development of teacher training programs to adjust to the new system and the provision of refresher courses for school administration in financial management, procurement, and upkeep of school premises. Similarly, we also saw that the reforms could have benefited further from being multi-dimensional. They needed to increase the capacity of teachers and to improve facilities in schools. This was especially the case in poor provinces that were marginalized and deprived of critical education material.

24. Conflict and political upheaval can present a window of opportunity to introduce major education reforms. The ANC took power at a time when the country was in a difficult position. But taking power also gave them a unique window of opportunity through which to introduce

[287] OECD (2008); Kraak and Young (2001); TIMSS (2015).

transformative reforms in the country. The government used this chance and harnessed their political capital to introduce quite transformative reforms that may have been more difficult to pass through at other points in time. This can serve as a lesson to policy makers who may be working in settings where there is a lack of stability or are recovering from disaster.

A silver lining may also be found in a new government's ability to use this time to make positive sweeping changes that otherwise might not have been possible. However, this makes sense only if the government then improves the human capital of the country to successfully execute reforms and make public the agents of change, as in the case of Rwanda or Singapore, for example. However, in the case of South Africa, the government often remained conflicted about what the priority should be: personal gain through power or equitable, quality education.

Country case 7: Implementing primary education reforms in Bangladesh, 2004–2015

Introduction

1. Bangladesh is known globally for its development achievements and potential, including meeting several of the UN Millennium Development Goals. In terms of primary education, Bangladesh's progress has been mixed. Primary school access has markedly improved, as has gender parity. But the country has yet to make sufficient progress in key areas of education quality. Primary dropout and repetition rates are high, the transition to secondary education is improving but still low, and learning outcomes have stagnated.

2. This case study draws from the lens of political economy to investigate why the country has not done better when it comes to improving quality. Over the last decade, Bangladesh has worked with its development partners to implement sector-wide education reform programs to improve quality. It recently completed the third iteration of a policy called the Primary Education Development Program (PEDP) (2011–2015). But various sample-based assessments starting in the early 2000s (e.g., Education Watch's learning achievement test) and the more recent National Student Assessments (NSAs) for grades 3 and 5 (implemented in 2011, 2013, and 2015) suggest that learning levels are still low.[288]

   For this case, we focus on the implementation experience of the second and third PEDP policies (i.e., 2004–2015) to shed light on how accountability to learning has been introduced through introducing decentralized school management programs at the local (*upazila*) level, as well as through examinations to track learning. The case argues that there are few incentives for the government to enforce measures at the local level to ensure effective learning. While the education system is decentralized, without an upward flow of accountability, and in the presence of rent-seeking through patronage and competitive clientism,[289] incentives to deliver quality education remain low.

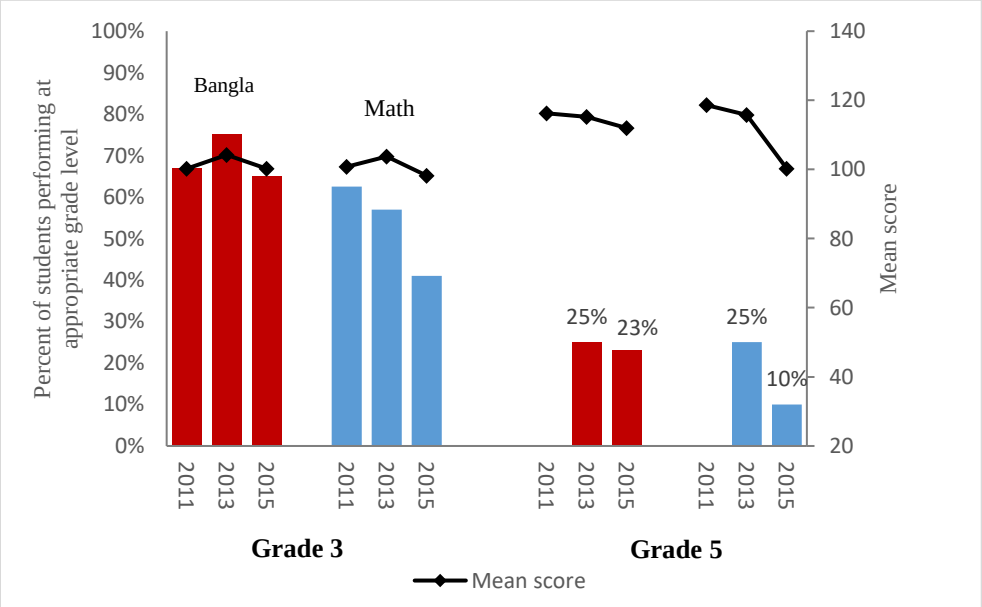
Data/evidence: Trends in learning outcomes and definition of reform episode/period

[288] World Bank (2013b).
[289] Hossain et al. (2017).

3. Primary level completion rates, transition rates, and learning outcomes remain low, particularly in some of the poorest regions of the country.[290] The NSA shows dismal levels and no improvement since 2011 in either third or fifth grade; 25% of fifth graders performed at their grade level in 2013 in Bangla,[291] and this number was estimated to be 23% in 2015.[292] A 2008 study reported that 80%–90% of students do not learn grade 5 level skills in basic competencies in math, reading, and writing until they are in secondary school.[293] And yet half of those students who go onto secondary school end up completing at all, with girls more likely to drop out than boys. Only 60% of children complete primary school on time and just 44% transition to secondary school. [294]

4. Learning outcomes can be best explained by the school students attend rather than by their individual backgrounds. The 2013 NSA suggested that school-related factors explained 73% of the differences in students' performance among grade 5 students, while just 27% of the differences could be explained by individual student factors.[295] This suggests that quality may have more to do with teachers, management, governance regimes, and schools than it does with the situation of students themselves,[296] what Hossain calls "frontline discretion."[297]

Figure A8: National Student Assessment[298]—Percent performing at their grade level



Source: World Development Report team using data from the World Bank (2013a).

[290] World Bank (2013a).
[291] MPME (2013).
[292] The 2015 estimates are from a draft World Bank report.
[293] World Bank (2013a).
[294] World Bank (2013a).
[295] World Bank (2013b).
[296] Hossain et al. (2017); World Bank (2013b).
[297] Hossain et al. (2017, p. 8); World Bank (2013b).
[298] World Bank (2013a).

Broader context

5. Following independence from Pakistan, the Awami League introduced a socialistic and secular program for social and economic development.[299] In terms of education, this resulted in nationalizing 26,000 community primary schools, where "education was declared compulsory, secular and modernized—the last a reference to madrassahs/Islamic schools."[300] After the mandate, Bangladesh entered a period of military rule for 15 years. During this time, the dominant elite coalitions supported madrassahs and bolstered links with unions.[301] They also decentralized public administration and made the state ideology Islamic. Both were popular moves that also influenced education policies.

6. As a way to consolidate power, the Awami League attempted to bring about political compromise through initiatives such as the 1990 Compulsory Primary Education Act, which established a free and compulsory education as national policy.[302] But it occurred too late and the regime was toppled later that year. Multi-party democracy returned in 1991, and the most recent election occurred in 2014, when the Awami League kept power in an uncontested election.

7. The emphasis placed on nationalization or Islamic education has changed depending on who is in power. But improving the formal education system has been a popular and consistent platform throughout. The government has consistently emphasized the need to expand access. To overcome structural barriers, it has introduced policies to build schools, hire teachers, and give conditional cash transfers to the poor.[303]

8. The Bangladesh education system caters to over 30 million students and involves many different actors. For example, there are 13 different types of providers in primary education, and 98% of secondary schools are private but receive public funding.[304] Administration of the primary education system (grades 1–5) is organized by the Ministry of Primary and Mass Education (MoPME), which is also responsible for nonformal education and literacy. The Ministry of Education oversees secondary education (grades 6–12), madrasah education, technical and vocational education and training, and higher education.[305] Members of parliament (MPs), local government officials, and education officers play a key role in school management.[306]

[299] Hossain et al. (2017).
[300] Hossain et al. (2017).
[301] Hossain et al. (2017).
[302] Hossain et al. (2017).
[303] Hossain et al. (2017).
[304] World Bank (2013b).
[305] World Bank (2013b).
[306] Richards and Vining (2016).

9. The government is the largest contributor to education financing in the country. Between 2000 and 2008, they spent between 2.2% and 2.5% of GDP on education. This figure is lower than other low incomes countries that spent at least 3.2% of GDP on education during the same period.[307] Ninety percent of public funds are used to cover teacher salaries in both government and nongovernmental schools, which leaves few resources for other education-related expanse. Further, the education system is compounded by "high levels of inefficiency"[308] that suggest more funding will not necessarily lead to improved learning outcomes due to "underutilization of funds, or misuse of funds."[309]

Education reform program

10. The government worked with development partners to improve primary education through the PEDP2 and PEDP3. They were separate policies but were designed to offer some continuity to the general trajectory of the education system with respect to quality. PEDP2 was introduced in 2004 and was considered an important milestone for strengthening primary education. For example, the budget committed nonsalary recurrent spending to produce and distribute free textbooks to primary school.[310]

    The PEDP2 also focused on coordinating with development partners, taking steps for revising the curriculum, improving teacher professional development though a diploma in education, providing examinations to all grade 5 primary students, and introducing learning assessments.[311] At the local level, PEDP2 sought to improve SBM through introducing school-level improvement plans as well as establishing monitoring and evaluation systems across the sector.[312]

11. PEDP3 was designed to build upon PEDP2. Carried out between 2011 and 2015, it aimed to "establish 'an efficient, inclusive, and equitable primary education system delivering effective and relevant child-friendly learning to all Bangladesh's children,' through six results areas: i) learning outcomes, ii) participation, iii) reducing disparities, iv) decentralization, v) effective use of budget allocations, and vi) program planning and management through activities in 29 sub-components."[313] PEDP3 also focused on expanding access and improving equity to the poorest regions and groups across the country.[314]

12. To improve quality and effectiveness, PEDP3 meant that more teachers were to be trained in subject content and pedagogy. In 2008, just 20% of teachers had training in the core curriculum, and 40% of female and 60% of male teachers had no training in the subjects they

[307] World Bank (2013b).
[308] World Bank (2013b, p. 38).
[309] World Bank (2013b, p. 38).
[310] World Bank (2013a).
[311] World Bank (2013a).
[312] World Bank (2013a).
[313] Hossain et al. (2017, p. 17).
[314] Cameron (2010); Hossain et al. (2017).

taught.[315] The government made a diploma in education program part of the certification process,[316] with opportunities for continuous professional development and training, and they created support networks for teachers. The government also introduced school and student assessments. The idea was to monitor school and student performance to provide accountability for learning outcomes.[317]

13. Another important change that came about during the reform period was the nationalization of private schools and teachers into the government's education system. This was not directly connected to the PEPD3 policy. Rather, the change came about through "longer-lasting agendas," including the ongoing concern private teachers had concerning disparities in pay and status.[318] As later sections of this case will describe, this move also strengthened the collective power of teachers as a political group that impacted the effectiveness of the PEDP reform.

Motivations/drivers of education reform

14. The funding reforms were mostly financed by the government with the support of development partners. They began the reforms to address low learning levels. Currently, there is currently some competition between elites through different party and political agendas. However, one commonality seems to be the sustained elite support for improved quality education, something that is important for "legitimacy, supporters, and policy success."[319]

15. There is little evidence to suggest that reforms were driven by the demand for skilled labor by the business industry. Low demand from skills means presumably little pressure from the elite to improve workforce skills, even though this may eventually come to be the case as the economy modernizes. On the other hand, teachers' interests have served as a powerful interest group due to their ability to organize as public sector workers to shape policy that benefits their situation.[320]

Reform process

Implementation

16. The government failed to fully implement PEDP2 because of the national education system's complex bureaucracy. For example, many of the PEDP2 reforms hinged on the success devolution of responsibilities and management to districts, upazilas, and schools. However,

[315] CAMPE (2009).
[316] World Bank (2013a).
[317] Hossain et al. (2017).
[318] Hossain et al. (2017, p. 28).
[319] Hossain et al. (2017).
[320] Hossain et al. (2017, p. 28).

oversight of these local levels of administration were beyond the jurisdiction of the MoPME.[321]

17. Existing evaluations and reports point to the ineffectiveness of teacher training and the lack of incentives for teachers to perform well. The PTIs have been described as lacking "organizational skill, teaching proficiency and capacity, as well as adequate physical infrastructure."[322] Teachers do not have enough training or incentives, while the curriculum and testing continue to prioritize and reward rote learning and memorization. Teachers have low motivation because the profession is still not well regarded.

    There are few incentives for innovative approaches to learning and teaching. These few incentives mean that additional years of teaching experience are not correlated with higher student learning outcomes.[323] Further, teachers have limited contact hours with students due to training, meetings, and tardiness. In addition, 90% of primary schools are double-shift schools, resulting in 20% fewer contact hours with students compared to single-shift schools.[324]

18. The recruitment and hiring of teachers were further impacted by informal politics. Group rivalries and disputes emerged during teacher recruitment. Teachers were not being recruited into placements based upon merit, knowledge, or needs but rather through political calculations and clientelism. Currently, teachers reportedly pay local education officers and politicians,[325] meaning teachers are not being placed according to merit or need but instead on their capacity to make payments to local officers. Richards and Vining[326] speculate that if teachers know their job security is linked to their ability to make a payment, they may be less likely to take their training seriously. In turn, teachers are willing to make this payment because they can expect to earn back this income from private tutoring.

    Richards and Vining also offer an account for how this form of corruption may serve to undermine the government's efforts to improve education quality in its entirety: "the informed judgment is that teachers' unions accede to illegal payment made by prospective teachers to MPs and officials, and do not protest this practice vigorously. In exchange, teachers' unions expect lax monitoring of teachers' classroom performance and acceptance of teachers undertaking private tutoring of their students." The government's motivation in introducing the PECE may in part have been to increase hours spent teaching without disrupting conventions about lax monitoring of teachers and private tutoring. Despite PECE's

[321] Hossain et al. (2017); World Bank (2013a).
[322] Hossain et al. (2017, p. 21).
[323] Hossain et al. (2017); World Bank (2013b).
[324] World Bank (2013b).
[325] Richards and Vining (2016).
[326] Richards and Vining (2016).

evident weaknesses as a test of student learning outcomes, teachers' unions have acquiesced in it because it has induced parents to demand additional private coaching.[327]

Information and knowledge

19. The PECE is a national examination system introduced in 2009. PECE is administered to grade 5 students across the country as a way of providing a uniform method to understand primary school completion, offering more accountability,[328] and providing a credential of primary school completion that could be used by both schools and employers. Decentralization may have had an effect on accountability measures and how data are used. For example, Richards and Vining note that "the assessment of primary school completion in government schools was decentralized to the upazila level and was not subject to a centralized operation that seems to have created a political incentive to inflate the completion rate."

    When the PECE was introduced, 95% of those who took the test passed.[329] The presence of the PECE is an important step toward monitoring and evaluating learning outcomes, but it is also subject to manipulation. When the test was introduced, cheating and memorizing answers was common, with answers reportedly available on the internet.[330]

20. The challenges inherent in a national examination are further exacerbated with teachers who also work as private tutors. Approximately 77% of households with children preparing for the PECE do so with the help of paid, private tutors.[331] The majority of tutors are teachers who moonlight for extra money. While the problem of leaking answers has been reduced, the PECE has a set pattern of questions used annually, which encourages one to memorize answers through tutoring.[332]

Coalitions and incentives

21. There are over a dozen teachers' unions in Bangladesh, consisting of about half a million members who exercise "considerable autonomy."[333] An example of their success has been nationalizing all registered nongovernment primary schools (RNGPS). It means that all teacher salaries, ranks, and status are part of the government systems. This activity occurred during PEDP3 but was not part of the policy. Aside from their primary role as educators, teachers have important political functions. For example, they serve as polling officers during general elections and help staff public health campaigns. However, the strongest sign of their political significance is that a further 344,000 teachers employed in the independent RNGPS have been made government employees.[334] This has increased the amount of teachers. They

[327] Richards and Vining (2016, p. 17)
[328] Richards and Vining (2016); Nath and Chowdhury (2009).
[329] Richards and Vining (2016).
[330] Richards and Vining (2016).
[331] CAMPE (2015).
[332] Hossain et al. (2017).
[333] Hossain et al. (2017, p. 11); Richards and Vining (2016).
[334] Hossain et al. (2017).

are now nearly half a million middle-class voters, "giving them considerable clout in the national policy space"[335] while also guarding against accountability based upon learning outcomes.

Innovation and agility

22. The education system is bureaucratic and complex. Decentralization has allowed for some flexibility to implement at the level of the upazila, but this has not been accompanied by an accountability to learning. PECE's introduction has been a step in the area of accountability, but this system was subject to manipulation by teachers. Examinations also limit the amount of creative teaching that can be done, as it has imposed limits for innovative approaches to teaching.[336]

23. Perhaps due to the size and complexity of the education sector, the government has worked closely with nongovernment providers of education. One example is BRAC, a large international NGO that started in Bangladesh. It provides nonformal education to 270,000 out-of-school children as a way to channel them back into formal schooling. This type of collaboration has existed in the current political settlement under the Awami League. In the past, governments had been reluctant to collaborate with or formally recognize NGOs as development partners.

24. Lack of incentives for good performance may also stunt innovation and agility. There are few incentives available to improve performance, and the interface between policy and practice and the incentives for strong performance is weak. This is despite the fact that the "regulatory regime is reasonably strong, demanding transparency, accountability, and good-quality performance."[337] However, the lack of incentives for compliance (or disincentives for noncompliance) hamper the potential effectiveness, regardless of how well formed the policy is.[338]

Lessons learned

25. Decentralization is a complex political process, from national to local levels. The size and complexity of an education that caters for 30 million young people and half a million teachers has inherent challenges. On one hand, a devolved system is vital to ensure the education system is responding to local needs. On the other hand, Bangladesh needs a sophisticated national education system capable of offering a national strategy that can be implemented. It must be simple, robust, and measurable.[339] Richards and Vining suggest one idea to improve the effectiveness of PEDP may be to pilot additional standards of

[335] Hossain et al. (2017, p. 16)
[336] World Bank (2013b).
[337] World Bank (2013b, p. 51).
[338] World Bank (2013b).
[339] World Bank (2013b).

accountability in a selection of upazila to see what kind of incentives and accountability measures could work best for bringing about improvements in learning.[340]

26. While national assessments have been conducted, their results have not impacted the education system as we might have otherwise expected. This can be explained, in part, through politics. Bangladesh has a highly centralized education system that paradoxically allows for local discretion when it comes to implementation.[341] Informal political alliances and bargaining allowed for perverse incentives when it came to the recruitment and hiring of teachers, the role of the PECE, and the ubiquity of private tutoring by teachers. This agency, on the part of local government education systems, contributed to some of the poor learning outcomes. However, evidence was also presented to suggest that the quality tended to vary school-by-school. Thus, it may be possible for the national and local education systems to introduce incentives to motivate teachers and local government officers to improve learning outcomes at the school level.

27. Teacher development and incentives were overridden by patronage and clientism. The unofficial incentives for teachers did not align with improved learning for the students. Like many developing (and developed) country contexts, the teaching profession was (and still is) stigmatized as one of low status and was compensated accordingly. The motivations for teachers to enter the profession were not aligned with the goals of the PEDP. Efforts to introduce additional forms of accountability were presumably blocked by powerful teachers' unions. This lesson learns that without the introducing policy measures to address teacher performance and motivation, quality is not likely to improve.

28. Efforts to improve education quality existed despite a lack of demand for skilled labor by the private industry. To date, the drive to improve public education has come from the government and its development partners (and, to a lesser extent, households who may remove their children from public schooling if the quality is poor). If and when Bangladesh requires a more educated workforce, and as the population itself slowly becomes more educated, the demand for better education, including a greater percentage of GDP, may come from these actors.[342]

Country case 8: Implementing teacher certification reform in Indonesia

Introduction

1. The established link between teacher quality and learning outcomes[343] suggests that investing in the teaching workforce should strengthen education quality. But Indonesia's experience

[340] Richards and Vining (2016).
[341] Hossain et al. (2017).
[342] Hossain et al. (2017).
[343] Chetty, Friedman, and Rockoff (2011).

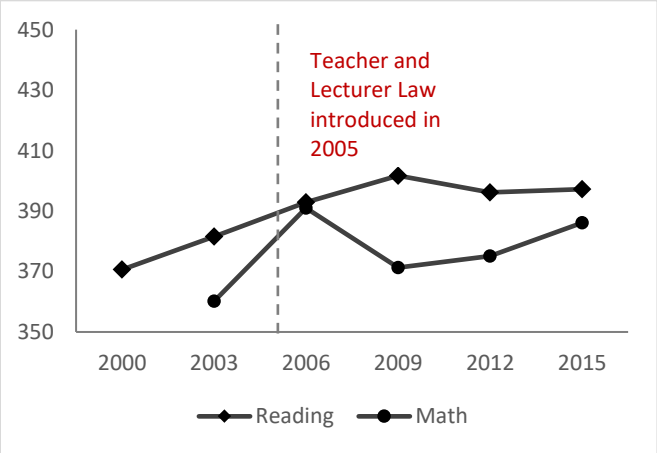
suggests that this outcome is not guaranteed. This case study investigates the implementation of Indonesia's Teacher and Lecturer Law 2005 (hereafter referred to as the Teacher Law), focusing on a component requiring teacher certification. Under this reform program, the government made massive investments in improving teacher quality. The government introduced a certification and financial incentive program that effectively doubled teacher pay.

However, capacity and political challenges hampered implementation. Introducing this reform coincided with mixed trends in learning outcomes. Ultimately, the Teacher Law turned out to be a drain on the public budget, and improvements in quality were not made. This case study describes the reform process and political conditions to help understand why a promising education reform did not lead to improved learning.

Data/evidence: Trends in learning outcomes and definition of reform episode/period

2. While Indonesia has made progress in expanding access, especially in primary,[344] learning outcomes have been mixed. Trends in international assessment scores around the 2005 Teacher Law show a mixed picture, indicating the reforms did not improve learning. While Indonesia experienced steady improvement in PISA reading performance between 2000 and 2009, performance in math declined.[345] The steady improvement in reading scores until 2009 cannot be attributed to the 2005 reforms since Indonesia was already improving at the same rate from 2000 to 2003 and up to 2006 (see Figure A9). In short, Indonesia's education quality has not improved as well as we might have expected given the government's investment in expanding and professionalizing the teaching workforce around 2005. The sections that follow turn to examine the political and economic dimensions that have impacted quality.

Figure A9: Mean scores in PISA reading and math over time



Source: OECD (2016).

[344] By 2005, gross enrollment rate in primary was above 100%. Enrollment in secondary was much lower, at 60%.

[345] Trends in performance in other international assessments also show an improvement in reading but a decline in math scores. While PIRLS primary reading scores increased between 2006 and 2011, math performance in TIMSS (secondary) declined by 11 points from 2007 to 2011.

Broader context: Overall political and economic context

3. Education policies during President Suharto's New Order in the 1970s to 1980s had lasting implications for teacher quality. A massive effort to expand education resulted in thousands of new primary schools between 1975 and 1987. New teachers were hired rapidly albeit with little pre-service or in-service training.[346] As of 2005, this teacher cohort had roughly another decade of work left before retirement. [347] This meant that subsequent administrations had to contend with low average teacher quality. Moreover, teacher management during the New Order suffered from bureaucratic complexities, and the school system was used as a means to extract rent and get votes.

4. The fall of the New Order in 1999, precipitated by the Asian financial crisis,[348] made way for reforms like decentralization in the early 2000s. By shifting responsibilities for teacher management from the central government to districts, decentralization aimed to reduce bureaucracy and improve efficiency. However, decentralization did little to streamline teacher management. Local governments had increased resources and political power, enabling them to continue patronage practices. Moreover, there were no tools to hold districts accountable in following a centralized policy framework such as the teacher certification law.[349]

5. New policies in the 2000s increased education spending, yet this increase evidently had little impact on learning outcomes. For example, the Education Law of 2003 mandated that the government allocate 20% of its expenditure on education.[350] The Bantuan Operasional Sekolah (BOS), introduced in 2005, transferred funds directly to schools to finance operational needs including hiring more teachers. And between 2000 and 2006, education spending doubled and continued to increase over time.[351]

6. Other economic and policy changes do not seem to explain mixed learning outcomes after the Teacher Law. Since 2000, Indonesia has enjoyed steady economic growth. Despite this, enrollment has declined. Between 2005 and 2007,[352] a shift in enrollment at the lower secondary level occurred that coincided with the drop in lower secondary math performance after 2006. However, reading outcomes improved despite higher enrollment.

7. Despite the economic growth, teacher salary in real terms had been declining. By 2005, teacher salary was lower relative to other professions requiring similar education levels. Low wages

[346] Chang et al. (2014).
[347] In 2005, close to 77% of Indonesia's teachers were 41 years of age or older and retired at age 60; see Chang et al. (2014).
[348] Rieffel (2007).
[349] Chang et al. (2014); Rosser and Fahmi (2016).
[350] This was done in part to fulfill a 2002 constitutional amendment mandating free basic education for all.
[351] Between 2007 and 2012, real per-student spending increased by 24% (34%) at the primary (secondary) level.
[352] The GER in lower secondary increased from 72% in 2005 to 86% in 2007. In subsequent years, GER remained above 80%.

were perceived to have reduced teacher motivation and also made the profession unattractive for qualified individuals.[353]

Education reform program

8. Spurred by low teacher statuses and disappointing learning outcomes, the Indonesian government introduced the Teacher and Lecturer Law in 2005. The law provided a framework for a comprehensive reform of teacher development and management. It covered all stages of the teacher development cycle from setting teacher standards, certifying teachers, and attracting qualified teachers to mandating continuous training for career advancement.

9. Teacher certification was a key reform component to address quality. Upon completing certification, teachers were eligible for an allowance that effectively doubled their base salary. However, teachers had to meet three requirements to be certified. First, they had to complete a four-year degree.[354] Second, they had to teach a minimum of 24 periods (roughly 18 hours) per week to maintain certification. Third, teachers had to undergo a competency assessment. The necessity for such an assessment and its form became a major bone of contention. While the Ministry of Education and Culture (MOEC) wished to assess competency through a written test as well as through classroom observations, teachers' unions did not think assessments were necessary at all. Ultimately, it was decided that teachers would either be evaluated on their teaching portfolio or on a 90-hour remedial training and subsequent assessment. Legislation required that all government school teachers get certified by 2015.[355]

10. Reform elements also included measures to improve quality and attract qualified candidates into the profession. The government introduced a six-month (one-year) long professional training course for primary (secondary) school teachers after completion of a four-year degree. The government also sought to improve pedagogy through better pre-service and in-service training. The reform included providing teacher support through improved processes of induction, probation, and mentor opportunities. To balance teacher distribution, the Teacher Law included a special allowance for those teaching in remote areas.

11. There were several positive elements that made the reform design seem promising at the outset. To guide the reforms, the government created a National Education Standards Board in 2004 that established minimum standard and processes for evaluation. A new Directorate General for Quality Improvement of Teachers and Education Personnel was also created within MOEC to oversee teacher management. By conditioning the allowance on certification instead of simply increasing salary for all teachers, the Teacher Law was designed to jointly address issues of quality and low teacher pay.

12. At the same time, key features of the reform were misaligned with learning and were incoherent with existing capacities, which may have contributed to its ultimate failure. Most notably, the Teacher Law gave a one-time permanent salary increase through the allowance but lacked

[353] Chang et al. (2014).
[354] Before the reform, teacher training universities offered certificates based on three- or four-year education programs. Further certification was not required by the Ministry; Chang et al. (2014).
[355] Chang et al. (2014); De Ree et al. (2015).

measures to hold teachers accountable for continued performance. Another shortcoming was that appropriate measures were not taken to ensure that the new directorate was adequately staffed to implement the reforms.

13. It is not surprising that the impact evaluations of these reforms have shown little positive change. Teacher salaries were found to have no effect on student outcomes between 2009 and 2012.[356] The policy's only positive impact was on increasing financial security and reducing the prevalence of second jobs. In addition, no differences were found in learning outcomes between students taught by certified teachers compared with uncertified teachers.[357]

Motivations/drivers of education reform

14. Although low teacher status and pay were the original triggers for the reform, policy makers also realized the need to address low teacher quality. An analysis by MOEC's new directorate suggested that poor pay had resulted in the prevalence of second jobs, high absenteeism, and low status. These factors dissuaded qualified candidates from entering the profession. The analysis also revealed that most teachers had been hired during Suharto's expansion program with little training and were producing low student outcomes. This realization convinced MOEC that higher status had to be met with better quality.

15. The evolving role of teachers due to decentralization and increased school funding also motivated policy makers to increase teacher capacity through implementing the Teacher Law. Decentralization had given schools more authority in improving quality. Moreover, increased direct funding for school operations through programs like BOS warranted greater teacher participation in improving school quality.

16. The primary motivation for teacher associations to support the reform was the prospect of higher status and pay. Greater freedom for teacher activism in the post-New Order era enabled teacher associations, especially the Indonesian Teachers Union,[358] to voice their demand for improving teacher income.[359] Moreover, the parliament consisted of many former teachers who had taught during the New Order and could relate to teacher concerns. These members recognized the potential of the teaching force as an important voting group and rallied their party leaders to support the reform.

Reform process

Implementation

17. While broad political support made the Teacher Law seem promising, implementing the reform, especially the certification component, was ineffective. The initial proposal for a competency test was met with stiff opposition from teachers' unions. Given that teachers were important constituents, the parliament supported their position and refused to fund competency tests. As a compromise, tests were replaced with an assessment of a teacher portfolio

[356] De Ree et al. (2015).
[357] Chang et al. (2014); De Ree et al. (2015).
[358] It claims to have a membership of more than 1.9 million teachers.
[359] Rosser and Fahmi (2016).

comprising personal references, publications, documentation of in-service training, and model lesson plans. Those who failed the portfolio assessment had to partake in a 90-hour remedial training and take an exam based on the training. Such watered-down certification requirements hampered any meaningful enhancement in quality. For example, the 90-hour training program had a 100% pass rate. Furthermore, there were reports of falsified teacher portfolios.

18. Implementation was also hampered by capacity issues. The new MOEC directorate was not staffed enough to oversee the teacher management process. The directorate was dismantled around 2008, creating further delays. Select education institutions were tasked with assessing teacher portfolios. However, the number of chosen institutions were not enough to cover a large number of portfolios. Moreover, delays in funding for these institutions further hastened the assessment process. Consequently, portfolios were not evaluated carefully.

19. Given the large number of teachers, and the fiscal impact of paying allowances, MOEC had adopted a staggered approach based on a systematic queue for entering the certification process.[360] However, local authorities disregarded these rules and selected teachers based on favoritism. Moreover, a set deadline for certifying all teachers pressured local governments into certifying as many teachers as possible without paying attention to quality. Local governments selectively hired teachers and used certification as a tool for rent extraction. Demand from the Ministry of Finance to spend the budget for professional allowance also incentivized MOEC to certify teachers quickly.

Information and knowledge

20. Although not an original trigger for the reforms, information on education quality, as evident from reports measuring teacher competency and student outcomes, was important. One of the first tasks of the new directorate was developing a teacher profile, the results of which helped trigger the reform. Information on learning outcomes, however, were not systematically deployed in designing the reform or in monitoring the implementation progress.

21. When proposing the reform, available information on learning outcomes from international assessments were also not articulated to teachers and other stakeholders. Media coverage of the reforms revolved around teachers' rights and welfare rather than low learning outcomes. The need to improve quality was also not systematically communicated to other actors, such as parents and community members, who may have supported the competency tests. Lack of effective communication on the need to improve learning, given the competing priorities of policy makers and teachers, may have been a missed chance in building a long-lasting coalition.

Coalitions and incentives

22. Although there was broad consensus across policy makers, teachers' unions, civil society organizations, and parliamentarians on introducing reforms to upgrade teacher status and pay, this consensus broke down during implementation. Given the need to raise status, there was also an agreement on introducing a "professional allowance" rather than a general salary increase. However, debate on the means to prove competency fractured the initial consensus.

[360] Permanent teachers with civil service contracts were given priority, followed by temporary teachers with government contracts; see Chang et al. (2014).

23. While MOEC was not successful in building a sustained coalition, teacher associations, on the other hand, were effective in organizing themselves and lobbying local politicians to support their position on certification requirements and on transfers and promotion issues. For example, they argued that teachers who already had the combination of a four-year degree and some teaching experience should qualify for certification. In addition to lobbying for diluted certification requirements, teacher associations organized to jointly challenge another competency test introduced in 2012 in the Supreme Court.[361] Teachers' unions have also traded access to government grants from local politicians in return for mobilizing votes.[362]

Innovation and agility

24. To some extent, the Indonesian government has adapted some reform components based on implementation experience. For example, in response to malpractices around the portfolio assessment and certification process, MOEC eliminated the portfolio assessment. Instead, it was replaced by a pre-test of competency to determine whether the teacher would be in the certification queue and a post-test to determine certification. However, opposition to these tests eventually led them to lower the pre-test's passing score to 30%.

25. Components of the 2005 Teacher Law were also deployed to address the inequitable distribution of qualified teachers. Political elites at the central and local levels tended to hire teachers based on demographic patterns to expand their voter base rather than on distributional needs. The introduction of a minimum requirement of teaching 24 periods per week to maintain certification compelled local districts to redistribute teachers. Since it was impossible for a subject teacher to teach 24 periods per week in the same school, local authorities in one district were incentivized to redistribute teachers across schools in rural and urban areas to help interested teachers meet this 24-period condition.[363]

26. MOEC also developed some credible measures to address teacher overhiring. For example, in 2011, officials placed a moratorium on new civil service appointments in regions where government spending on salaries exceeded 50% of the local budget. These regions were hiring cheaper and less qualified part-time teachers than full-time civil servants, who would then be eligible for higher salaries. The moratorium was deemed credible partly because it had implications for fiscal constraints.[364]

27. To meet the demand for a four-year degree, Indonesia has also experimented with distance learning. Hybrid Learning for Indonesian Teachers, a program under World Bank Bermutu project in 16 provinces, and the existing Open University provided learning materials to teachers in printed form throughout Indonesia. In 2009 alone, over 485,000 teachers were enrolled in Open University for upgrading to a four-year degree.[365]

---

[361] Teacher organizations led by the Indonesian Federation of Teacher Unions attempted to challenge the new competency test through a judicial review request in the Supreme Court; see Rosser and Fahmi (2016).
[362] Rosser and Fahmi (2016).
[363] Rosser and Fahmi (2016).
[364] Rosser and Fahmi (2016).
[365] Chang et al. (2014).

Lessons learned

28. Indonesia's experience in implementing the certification component of the Teacher Law offers valuable lessons for other countries attempting to improve teacher management and quality.

29. Governments should be mindful of political economy factors at all stages of the reform process. Even though there was broad support for reforms to address teacher welfare, different motivations for policy makers, versus teachers combined with teachers' political strength, weakened effective implementation. Considering the low accountability culture among most teachers hired during the New Order and their connections in parliament, MOEC would have benefited from adopting preemptive strategies to address potential political setbacks in implementation. For example, MOEC could have expanded its support base for competency tests through improved communication on learning outcomes.

30. Testing some options before committing resources to a blanket increase in payment may be prudent. A one-time increase in salary for all government school teachers did not guarantee an improvement in teacher quality. Instead, the government simply ended up shouldering a fiscal burden that is difficult to reverse. Starting with a targeted approach to improve quality may have reduced the immediate fiscal burden and avoided implementation capacity issues. For example, limiting incentives and certification requirements for new hires instead of teachers close to retirement, and gradually phasing in these reforms, may have been met with less resistance from the more politically powerful senior teachers.

31. Ensure accountability mechanisms are in place before investing in a nationwide financial incentive program for all teachers. This could take two forms. First, accountability measures must be continuous rather than a permanent one-time increase in salary. Second, given Indonesia's decentralized governance structure, tools must be developed to hold local governments accountable to follow decrees from the central government.

32. Ensure adequate capacity before implementing any reform. Understaffing in MOEC's new directorate, as well as the institutions for assessing teacher portfolios, added to implementation challenges. It is imperative to ensure relevant agencies are well aligned toward meeting reform objectives.

33. Elements of a reform program can be effectively deployed to address related issues. A good example of using a single policy to target multiple issues was the Indonesian government's mandate of the 24-period per-week requirement for certification, used to mitigate teacher distribution issues. With this unrelated policy for teacher distribution, the central government was able to bypass local politics to make teacher deployment more efficient.

Country case 9: Implementing accountability reforms in Mexico over two decades

Introduction

1. Since 1990, Mexico has witnessed three major education reforms to address accountability and improve learning outcomes for all. Reforms have ranged from decentralizing education to evaluating teacher competency to performance-based pay, yet progress on learning outcomes has been mixed. A recurrent feature across the reform episodes has been the uneven power dynamics between Mexico's formidable teachers' unions and education authorities. In 1992, the central government, under President Carlos Salinas, state governors, and the teachers' union, signed an agreement that decentralized system operations and introduced a merit-based pay program. The early 2000s included another agreement with a new administration to improve education quality.

   Yet another alliance with the teachers' union was created in 2008 to amend teacher policies. However, misalignments in reform design and failure to address political constraints hampered implementation. In 2013, the Pena Administration adopted a more confrontational approach and introduced a reform package to overhaul the teacher hiring and evaluation system. In the absence of broad-based consultations and effective communication, these reforms were met with stiff resistance and have yet to be fully implemented. This case outlines the characteristics of these reforms and of the reform process with a focus on the 1992 reforms and the most recent reforms proposed in 2013.

Data/evidence: Trends in learning outcomes and definition of reform episode/period

2. Despite attempts to improve education quality during each of the four administrations since the early 1990s, learning outcomes from 2000 to 2015 have not improved consistently.[366] Average performance in the 2015 PISA reading test was the same as the score in 2000. Although average reading scores among 15-year-olds increased from 2003 to 2009, this change represents a catch-up to the score in 2000 (see Figure A10). This suggests that there was no real change in reading performance over a 15-year period. Stagnant reading scores among sixth-grade students in the regional LLECE tests between 2006 and 2013 corroborate this trend.[367]

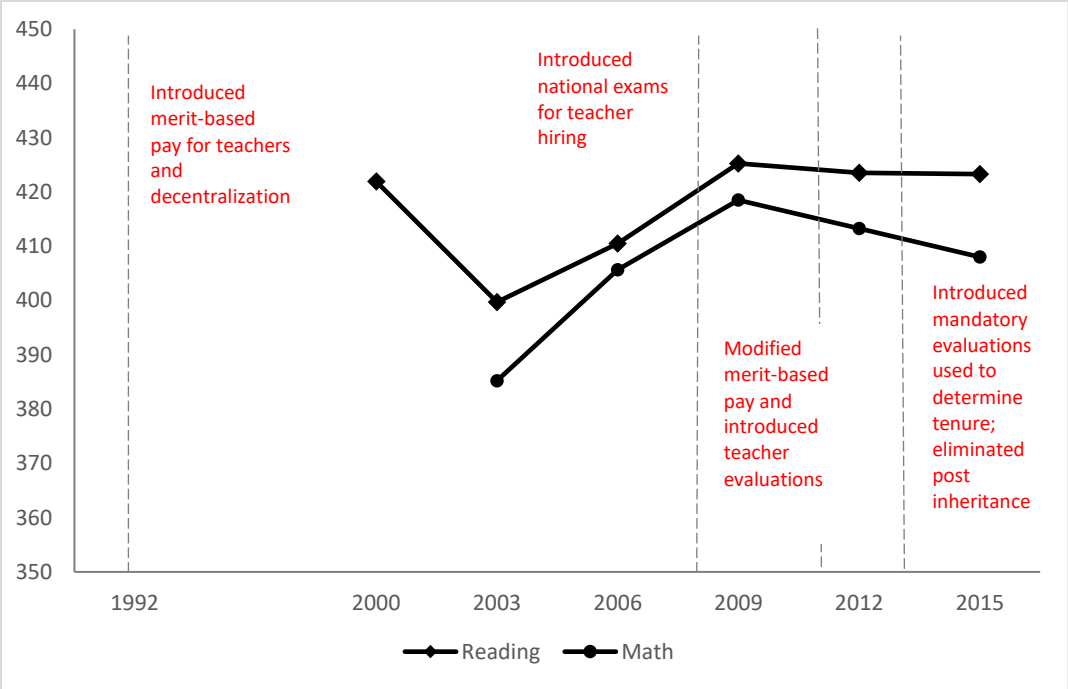
   Results from the PISA reading tests show that the gap between the topmost and poorest performers has narrowed from 2009 to 2015.[368] However, this decrease partly occurred due to a decline in performance among the highest scoring students. Performance in math has been more mixed; average scores in the PISA math tests jumped from 385 in 2003 to 418 in 2009. Over this period, Mexico also experienced a decrease in the share of low math performers. However, scores began to decline steadily in subsequent rounds in 2012 and 2015.

3. It is unclear whether trends in other education factors such as enrollment might have impacted learning outcomes. Since the early 1990s, Mexico has made significant improvements in gross and net enrollment at the secondary level. In particular, between 2000 and 2003, the enrollment rate among 15-year-olds jumped by 6 pp, which coincides with the drop in PISA reading scores

[366] The earliest information on learning is from the 2000 PISA tests. Mexico did not conduct learning assessments or participate in cross-national assessments before 2000.
[367] The average score in reading among Mexican sixth graders was the same (529) in 2006 as well as in 2013.
[368] OECD (2015). Tenth and 90th percentile scores in reading are available only for 2009, 2012, and 2015.

during this period.[369] However, despite relatively stagnant net enrollment at secondary level from 2009 to 2014, average scores in PISA remained unchanged (reading) or declined (math) after 2009. Enrollment at the primary level has been relatively higher since the early 1990s.[370] Following a trend of steady, but small, declines in primary enrollment since 1992, gross and net enrollment increased by around 2 pp between 2006 and 2013. Whether this small change in enrollment had any effects on primary level learning outcomes is unclear.

Figure A10: Mean scores in PISA reading and math over time



Source: OECD (2016).

4. The decline in PISA reading scores in from 2000 to 2003 coincided with the trend in economic growth around that time. GDP growth declined by 0.6% in 2001 but started steadily increasing until 2006. This trend may explain the eventual catch-up in scores to the 2000 level. Growth remained positive for subsequent years, except for 2009, when the financial crisis caused GDP to fall by nearly 5%. However, the economy recovered by the following year, and GDP growth remained positive in the next five years, averaging around 2.7%.[371]

Broader context: Overall political and economic context

5. A core feature of the Mexican education system is a powerful teachers' union that has deftly maintained its stronghold over education policy across different political administrations and

[369] From 2000 to 2009, enrollment among 15-year-olds increased from 52% to 66%.
[370] In 1992, gross enrollment at the primary level was 110%, while net enrollment stood at 97%.
[371] World Bank (2017).

despite factionalism within the union. The Sindicato Nacional de Trabajadores de la Educación (SNTE) enjoyed a longstanding corporatist relationship with the dominant political party—the Institutional Revolutionary Party (PRI)—given the appeal of support from a large voting group. Despite PRI's waning power with increased democratization in the late 1980s, SNTE's ability to mobilize votes during competitive elections enabled it to continue its influence even during the National Action Party (PAN) administrations from 2000 to 2012. The union inserted itself in key education positions. For example, education bodies were often composed of "mixed commissions" with equal representation from the Secretariat of Public Education (SEP) and SNTE.

6. Although previous proposals to decentralize education were ineffective partly due to union resistance,[372] a combination of political and economic factors in the late 1980s and early 1990s set the stage for limited decentralization through the National Agreement for the Modernization of Basic and Normal Education (ANMEB). President Salinas orchestrated a leadership change in SNTE, buying support from the new leader, Elba Esther Gordillo. Higher economic growth in the late 1980s after the Mexican debt crisis, and the introduction of a social fund program (PROSANOL), also helped Salinas's position.[373] Simultaneously, an alarmingly narrow presidential victory and impending congressional elections in 1991 meant that Salinas had to tread lightly.

7. Challenges faced in reform efforts are also partly rooted in education policies dating back to the years following the Mexican Revolution. The 1917 Constitution assigned responsibility for primary and secondary education to municipalities and states, respectively. To counter high illiteracy rates[374] and expand equitable coverage, the federal government expanded its role in education provision, starting with the creation of SEP in 1921.[375] This effectively created a dual system in most states, with schools run by two different levels of government. This dual system persisted in many states well until 1992,[376] when decentralization was introduced, creating challenges in implementation.[377]

8. The union's longstanding reach in education administration and political influence has fostered a system of patronage and low accountability. Well until the late 2000s, teachers were recruited by state-level committees with heavy union influence without a clear process. Furthermore, there is a market for selling teaching posts, while some posts are inherited most notably in states like Oaxaca and Chiapas, where the union's radical faction, Coordinadora Nacional de Trabajadores de la Educación (CNTE), have a strong presence. The prevalence of "ghost"

[372] For example, in 1982, President Miguel de la Madrid announced decentralization reforms but did not succeed; see Ornelas (2000).
[373] GDP growth rate in 1989 was 3.4%—up from 1.2% in 1988—and it continued to increase in the next couple of years; see Grindle (2004).
[374] According to the 1921 census, the illiteracy rate was 66%, an improvement of only 6 pp from 1910; see Gómez-Zaldívar (2014).
[375] Grindle (2004); Murillo (1999).
[376] For example, in the 1991–1992 school year, over 40% of basic education enrollment was in state-run schools in Mexico and Nuevo Leon states, whereas in poorer states like Oaxaca and Hidalgo, 99% of students attended federally operated schools; see Grindle (2002).
[377] Grindle (2004); Fierro and Velasco (2002).

teachers eventually lead to misuse of public spending. Currently, absenteeism rates are high, and an effective mechanism to fire underperforming teachers does not exist.[378]

9. While the Pena Administration adopted a more confrontational approach to overhaul teacher recruitment and improve accountability amidst growing public awareness, CNTE's continuing influence and broader economic factors may add to implementation challenges. The availability of learning information and efforts by nongovernment groups such as Mexicanos Primeros in publicizing state-level quality indicators, as well as union disruptions across the country, have heightened public awareness of the need for reforms. At the same time, support from new actors such as the left-wing party, the National Regeneration Movement (MORENA) have strengthened CNTE's opposition. Moreover, dismal approval ratings for President Enrique Pena following rising gas prices might mean stronger organized opposition from CNTE as well as other discontent groups.[379]

Education reform program

10. Over the last 25 years, Mexico has witnessed several attempts to raise education quality through decentralization and efforts to improve teacher hiring and accountability. The 1992 ANMEB decentralized school operations to the state level and created a merit-based teacher pay program. However, these policies did not achieve intended results. Subsequent PAN administrations during 2000 to 2012 proposed amendments to improve teacher accountability with limited success, and they also introduced a census-based assessment. However, the most radical reforms to overhaul teacher recruitment practices, and to an extent, reverse decentralization, were undertaken starting in 2013, once PRI was back in power after a long hiatus.

11. In 1992, ANMEB was introduced to the country, and they transferred responsibility for two-thirds of schools from federal to state authorities. However, SEP handled key functions of planning and policy making, annual wage negotiations continued to be conducted at the central level, and education financing continued to come largely from the federal government.[380] State authorities were responsible for hiring teachers, determining fringe benefits for teachers, and participating in more routine affairs.[381]

    ANMEB also introduced Carrera Magisterial (CM), a voluntary merit-based pay program whereby teachers were eligible to receive significant bonuses after participating in a year-long assessment.[382] Before the introduction of CM, salary was determined by seniority and education. Under CM, scores were assigned based on six factors—seniority, formal academic degrees, training courses, teacher tests, student performance,[383] and a supervisor evaluation.[384]

[378] Bruns and Luque (2014).
[379] Prensa Latina (2017); Martin (2017).
[380] Gershberg (1999).
[381] Ornelas (2006).
[382] Bonuses could be as high as over 200% of basic salary; Ornelas (2004).
[383] Until 2011, student performance was determined by classroom-based tests that made up 20% of the overall score; see Hecock (2014); McEwan and Santibáñez (2005).
[384] Hecock (2014); McEwan and Santibáñez (2005).

Final evaluations were done at the national level, while the decision to follow through with national recommendations for bonuses was left to the states.

12. Attempts to modify CM components and improve accountability in general were made over the next two decades. In 2002, President Vicente Fox and the SNTE leadership signed the Agreement for Quality of Education. The Fox Administration created the National Institute for Education Evaluation (INEE) to evaluate and to publicize education results. However, INEE was created as a SEP body without autonomy.[385] A census-based assessment, ENLACE, was introduced for the first time in 2006.

13. Another agreement between SEP and SNTE was signed during the Calderon Administration. The 2008 Alliance for Quality of Education, which ran until 2011, proposed modifications to teacher incentives and recruitment.[386] Reforms included replacing discretionary teacher recruitment by states with a more clear process based on teacher examinations. After three-year-long negotiations, the Alliance for Quality of Education also agreed upon changes to CM. They reduced the weight placed on academic degrees and removed the seniority criterion. Student assessments were given higher weight, turning ENLACE into a high-stakes test. Reform proposal also included mandatory teacher evaluation tests to be taken every three years.

14. While the proposed reforms under the Alliance faced implementation challenges mostly due to union resistance and the fragility surrounding impending presidential elections in 2012, the new PRI administration pursued a more radical reform agenda. In 2013, the administration introduced the Education Reform Bill. Reforms included plans to replace CM with a more technically sound merit-pay program to be managed by an independent INEE. Newly recruited teachers were required to undergo an induction period that included annual evaluations for the first two years before granting tenure. Teacher evaluations were made mandatory every four years; those who failed three consecutive evaluations were to be removed from their posts.

15. Limited success in implementing meaningful accountability reforms could be attributed to program characteristics that were misaligned with improving learning. An evaluation of CM by Santibanez et al. (2007) concluded that most items in the teachers' subject-matter tests were inadequate and demanded low-level cognitive responses.[387] Even after more than a decade since its introduction, there were no appropriate mechanisms to check the validity of teacher test results. Other determinants of the CM score that had no plausible link with improving learning, such as seniority, were not changed until 2011.

    Bonuses awarded through CM are permanent throughout the teachers' tenure. This feature undermines accountability—a key goal of performance incentive programs. It is no surprise that evaluations of the program have documented mixed results on learning. While Lopez-Acevedo (2004) finds that participation in CM is correlated with higher student scores, the

[385] During the Zedillo and Fox administrations, other education reforms such as the Pilot School project and the Quality Schools Program began. However, coverage of these school-based management reforms was not national.
[386] Bruns and Luque (2014).
[387] Santibañez and Rabling (2007).

degree of promotion has had no effect on student scores. Similarly, McEwan and Santibanez (2005) show that CM teachers facing stronger incentives did not improve test scores.[388]

16. Some features of decentralization were incoherent with one another and with states' existing capacity to assume full responsibility. Centralized wage negotiations meant that states had little discretion in managing resources given the large share of the education budget committed to salaries. Moreover, states lacked enough money to raise their own resources, as state tax revenues were low. Although the federal government introduced a compensatory financing policy to ensure equity across regions, some states had low discretionary resources. An evaluation by Gomez-Zaldivar failed to attribute improvements in coverage to the ANMEB's decentralization component.[389]

17. While administrations like those of Salinas and Pena took steps to help some political buy-in at higher levels, shortcomings in the reform design and process allowed political misalignments to persist. Salinas seized the opportunity to change SNTE leadership during growing internal conflict within the union. However, in exchange for union support for ANMEB, the Salinas Administration granted major concessions to SNTE, which further fostered clientelist practices in education. In addition to the across-the-board salary increase for teachers, a separate merit-based pay track was created for nonteaching staff—a feature that presented opportunities for patronage. The program was managed by a commission that included SNTE representatives, which resulted in conflicts of interest.[390] Moreover, the Fox Administration created the INEE, to evaluate teachers and students, as a nonautonomous unit within the SEP. This placed the INEE under SNTE's influence until the 2013 reforms.

18. The ANMEB's framework of decentralized operations with centralized wage negotiations exacerbated misalignments by creating an information asymmetry between education authorities and the union. SNTE exploited this asymmetry by winning high salaries and benefits for its members. For example, in negotiating fringe benefits at the local level, local union chapters received advice from central leadership, while state authorities, potentially from different political parties, lacked complete information on wage agreements reached at the national level.[391]

19. Although more recent reforms since 2013 include components that are likely to make the proposed teacher evaluations less unpopular, such as multiple improvement opportunities for poor performing teachers, the overall political strategy has not been expedient enough to counter the high level of opposition faced. The Pena Administration created a coalition of major political parties at the beginning of the administration. SNTE's leadership was also weakened with Gordillo's arrest under embezzlement charges. However, efforts to garner political support from state-level authorities were missing. An effective communication strategy to highlight the more benign features of the proposed reforms would also have been useful in countering misinformation regarding firing policies propagated by CNTE members.

[388] Santibañez and Rabling (2007) also do not find any impact on primary level scores; the study finds limited effect on secondary scores for teachers attempting to enroll in CM.
[389] Gómez-Zaldívar (2014).
[390] Ornelas (2004).
[391] Bruns and Luque (2014).

Strategic communication may have also been especially important given that the reforms were perceived more as a labor reform than an education reform.[392]

Motivations/drivers of education reform

20. Multiple hypotheses for the motivating factors behind decentralization reforms have been presented. Stated motivations behind decentralization and CM relate to improving efficiency and reducing inequity in coverage across states.[393] After years of centralized administration, increasing bureaucratic complexities and rigidities had made the system inefficient,[394] and there were disparities in illiteracy rates across states in the 1990s also motivated the ANMEB reforms. For example, in 1991 the illiteracy rate in Nuevo Leon was 4.3%, whereas in Chiapas, it was as high as 28%.[395]

21. Political factors may have also motivated the ANMEB reforms.[396] A potential motive for the PRI administration behind decentralization was weakening the SNTE, relative to the PRI, by forcing it to negotiate with 31 state governors instead of a single federal authority.[397] Another motive for decentralization could be limiting the growth of the dissident faction, CNTE, to a national-level movement and confining their presence to a few states. Gershberg (1999) argues that decentralization could also have been a way for the central government to absolve itself of responsibilities. While SNTE was wary of decentralization, favorable terms of the merit-based pay program, such as permanent pay increases, helped win their support. In addition, SNTE could also have used the program as a disciplinary tool, especially in states with high CNTE membership.[398]

22. After the early 2000s, growing public awareness became more important in driving reforms. For example, the INEE was established in response to public outcry after the delayed reporting of Mexico's results in the 1995 TIMSS test. The ENLACE tests also provided information on learning. Moreover, influential groups such as the Mexicanos Primeros helped communicate the need for quality-enhancing reforms. For example, the group produced the 2011 documentary *De Panzazo*, which highlighted the poor quality of Mexican schools.

Reform process

Implementation

23. The implementation of the various reform proposals over the past two decades has not been smooth partly because of pre-existing characteristics, such as the parallel education system, and misalignments in reform design. Although formal decentralization was put into effect immediately, some states faced significant hurdles. For states with two systems assuming responsibility for operating all public schools meant integrating two separate administrative

[392] Canedo (2016).
[393] Gómez-Zaldívar (2014); Ornelas (2004).
[394] Zorrilla Fierro (2002).
[395] Gómez-Zaldívar (2014).
[396] Ornelas (2004); Zorrilla Fierro (2002); Gershberg (1999).
[397] Ornelas (2004); Gershberg (1999).
[398] Hecock (2014).

bodies and systems. Conflict emerged between key actors, such as state-level ministry officials and institutes, to support transition to a decentralized system.[399] Insufficient capacity and resources also posed challenges in Tlaxcala and Oaxaca. Governors from these states returned education responsibility to the federal government in the early 2000s, citing insufficient resources. Ambiguity at the local level regarding federal and state responsibilities also reduced accountability.[400]

24. The implementation of teacher accountability reforms was also uneven. While states like Puebla and Guanajuato implemented CM without disputes, the program was delayed in Michoacan partly because of opposition from the dissident union faction, CNTE, which viewed the program with suspicion.[401] Indeed, SNTE used the merit-based pay as a tool to discipline or reward union members.[402] Recommendations from the central level based on CM scores did not always correlate with promotions and bonuses awarded by state authorities.[403]

25. Teacher recruitment reforms under the 2008 Alliance faced noncompliance from several states. Local union leaders in some states such as Nayarit and Guerrero persuaded authorities to continue discretionary hiring for state-financed teachers and to "uphold" inheritance rights. Moreover, standards for the teacher examinations had to be lowered because the majority of applicants scored below the minimum threshold. Furthermore, union leaders in Michoacan and Oaxaca disrupted the implementation of the teacher examinations as well as the ENLACE tests.[404]

26. The agreement reached in 2011 to introduce mandatory evaluations was not fully implemented, as the alliance with SNTE began to break down after strong criticism from union members. After a delay due to impending presidential elections, the first round of evaluations was conducted in July 2012. However, participation was low at only 53%. Reports of exams leaked online also surfaced in the media.[405]

27. Reforms proposed under the Pena Administration have faced massive resistance. Adding to implementation challenges were false rumors that teachers who had already received their posts through inheritance or illegal sales would be fired. School closures and protests by CNTE in the poorest performing states like Oaxaca, Guerrero, and Michoacan resulted in a loss of nearly 85 days of the school year in 2015. And support to CNTE by the newly formed MORENA created unexpected roadblocks.[406] Although states were supposed to have

[399] The experience of Guanajuato is an example, where a new institute to support the decentralization process was created. However, conflict over the control of education between this institute and the existing state-level ministry persisted for 18 months; see Grindle (2004).
[400] Grindle (2004).
[401] Hecock (2014).
[402] For example, dissenting members were excluded from participating in the program; see Bruns and Luque (2014); Hecock (2014).
[403] McEwan and Santibáñez (2005).
[404] Bruns and Luque (2014).
[405] Bruns and Luque (2014).
[406] SEP had cut funding to the Oaxaca Institute of Education, an important funding source for CNTE leaders. However, the union started getting federal funds from MORENA (all political parties receive federal money).

harmonized their laws in accordance with federal regulations by March 2014, a study found that at least 13 states had not complied with establishing sanctions for absenteeism, continued inheritance, and the existence of ghost teachers.[407] Despite a confrontational approach, the administration's sensitivity to mid-term elections resulted in a temporary suspension of teacher evaluations in May 2015. Although evaluations were reinstated, participation remains low.[408]

Information and knowledge

28. The role of information has become increasingly important over successive reforms. Systematic information on inputs and quality was not readily available when the ANMEB was devised. Until recently, SEP did not have reliable and systematic records of basic figures like the number of teachers employed. Although Mexico had participated in the 1995 TIMSS and performed poorly, results were not made public until 2001. The delayed TIMSS results created pressure on authorities to pay attention to quality and motivated the creation of INEE to conduct teacher and student assessments.

29. The introduction of the ENLACE tests in 2006 provided standardized information on learning to help guide policies. However, proposed reforms under the 2008 Alliance for Quality of Education did not align use of the ENLACE results with improving learning. ENLACE results were incorporated into CM, but its weight was increased from 20% to 50% of the total score, turning the assessment into a high-stakes test instead of a diagnostic tool for measuring learning.[409] This move was not only politically unpopular but also resulted in teaching to the test.

30. The Pena Administration built upon the previous administrations' efforts to construct a national census of schools, teachers, and students called the Sistema Nacional de Informacion y Gestion Educativa.[410] The census revealed the existence of ghost teachers, which motivated the decision to recentralize the teachers' payroll. Since January 2015, SEP and the Secretary of Finance have been in charge of payroll.[411] Moreover, nongovernment groups such as Mexicanos Primeros have made concerted efforts to spread information on the poor state of teaching and performances across the country. The group also administers their own assessments and publicizes state-level performances. In addition, increasing availability of information on learning and union abuses have increased public support for quality reforms.

Coalitions and incentives

31. Administrations across different reform efforts have formed coalitions; however, they have largely been limited within the upper echelons of select groups. The ANMEB reform was designed in consultation with SNTE leadership, yet political buy-in of rank-and-file members, especially those from the dissident faction, was missing. Although state governors also signed the agreement, they were not part of the negotiation or the design process. Most governors were reluctant to take on the administrative and fiscal burden that decentralization implied.

[407] Ornelas, Calderon, and Blikstein (2014).
[408] Canedo (2016).
[409] OECD (2013).
[410] Bruns and Luque (2014); Ornelas, Calderon, and Blikstein (2014); OECD (2013).
[411] Ornelas, Calderon, and Blikstein (2014).

However, wary of Salinas's tendency to remove recalcitrant governors, they signed the agreement.[412]

32. As in the case of ANMEB, the Pena Administration mostly took a top-down approach to reforms. On the second day of his administration, Pena formed a coalition with factions within the PRI and with leaders of major political parties, including traditional rivals such as PAN and the Party of the Democratic Revolution (PRD). The coalition created the "Pact for Mexico," and one of the first initiatives was education reform. However, there were no incentives for state-level administrators, who were (and are) key agents for implementing the reforms.

33. In recent years, SEP has taken some positive steps toward establishing a direct link with educators. For example, in 2014 SEP organized consultation forums consisting of teachers, principals, and supervisors to review the educational model. Three national and 18 regional forums took place between February and June 2014.[413]

Innovation and agility

34. Political constraints across administrations left limited space to pursue innovative reforms or to respond swiftly to shortcomings in existing reforms. For example, it took nearly two decades to introduce amendments to the teacher incentive program such as removing the seniority criterion and reducing the weight of academic degrees. Even then, the agreement reached after three-year-long negotiations was not fully implemented.

35. Nevertheless, some states managed to implement innovative reforms. For example, the state of Colima used the ENLACE tests to identify and to support poor performing primary schools. The Programa de Atencion Especifica para la Merjora del Logro Educativo (PAE), introduced in 2009, assigned technical advisors to diagnose test results and to help develop a school improvement plan.[414] Although this program was discontinued in 2011, it shows the possibility of using learning information for diagnostic purposes at the state level. Post-decentralization, Aguascalientes also took initiatives such as conducting a diagnosis of basic education efficiency and inviting parents, officials, teachers, and academics to participate in forums.[415]

36. In a departure from previous administrations, Pena took some innovative steps that helped codify reforms into law. The endorsement of the Pact for Mexico coalition helped legislate reforms, such as granting autonomy to INEE, in fast-track mode.[416] Unlike his predecessors, Pena did not created a pact with SNTE leadership that helped avoid severe watering down of reforms. To stop misuse of funds, the administration also eliminated the Federal Fund for Basic Education, which had been used to transfer resources to states for payroll and was abused by the SNTE.[417] While these strategies have helped pass reforms and reduced means for fund misuse, they have not proven to be useful for implementation.

---

[412] Grindle (2004).
[413] Chisholm and Leyendecker (2008).
[414] Patrinos (2015).
[415] Grindle (2004).
[416] Ornelas, Calderon, and Blikstein (2014).
[417] Ornelas, Calderon, and Blikstein (2014).

Lessons learned

37. Despite several reform proposals to improve quality and accountability over the last 25 years, misalignments in reform design and implementation challenges have not resulted in consistent improvements in learning outcomes. Mexico's reform experience offers valuable lessons for other countries, especially those with influential and well-organized opposition groups.

38. Ensuring reforms are well suited for the political environment is important for implementing meaningful reforms. Although merit-based pay can encourage better performance, the single teachers' union's longstanding influence at all levels of administration created an unfavorable environment for effective implementation. Not surprisingly, the voluntary program was not only uncorrelated with learning outcomes, but it also increased costs through the merit-pay component for nonteaching personnel, and it ended up being used as a tool to discipline union members. Mexico's experience suggests that accounting for such risks before introducing a performance incentive program is important. Starting with school-based incentives before launching individual-level incentives, as Chile did in the late 1990s, may also have been an alternative.

39. Lack of coherence in reform components can undermine effectiveness in the shorter and longer term. Assigning operational tasks and labor relations to the state level while keeping authority for key decisions at the central level limited the scope for effectively decentralizing the government. Furthermore, dividing salary and benefits negotiations at two different levels created asymmetry in information between union members and education authorities, which gave the union an advantage during subsequent negotiations.

40. Even when reform components are technically or politically aligned, an effective communication strategy is essential. Although the teacher recruitment and evaluation reforms proposed in 2013 adopted a carrots rather than sticks approach by allowing teachers who failed evaluations opportunities to improve before being fired, SEP's communication strategy was not effective enough to counter misinformation spread by opposition groups.

    Also, CNTE and its supporters often use a pro-democracy and anti-neoliberalism rhetoric to rally support. Therefore, to avoid negative press, highlighting reform elements perceived in a relatively positive light may be even more crucial to buy support in the context of a confrontational approach.

41. Strategic and broad-based consultations could help implementation efforts. Mexico's experience with decentralization and merit-based pay reforms designed in consultation with SNTE leadership shows how powerful groups can hijack policies to their advantage. Thus, building coalitions can be helpful, but doing so exclusively with groups holding ulterior motives may be detrimental. Moreover, coalitions encompassing different administrative and leadership levels may be more effective.

    Mexico's reform process has largely been top-down. However, as evidenced by Mexico's experience in implementing teacher recruitment reforms around 2008, and more recently in

2014, authorities in states may have mutually beneficial relations with union leadership. These authorities also require strong incentives to be willing to risk their political standing. While coalitions at the top, such as the Pact for Mexico, may help in passing reforms, creating incentives for local authorities, particularly in a decentralized structure, is likely to be crucial for sustained implementation.