Education

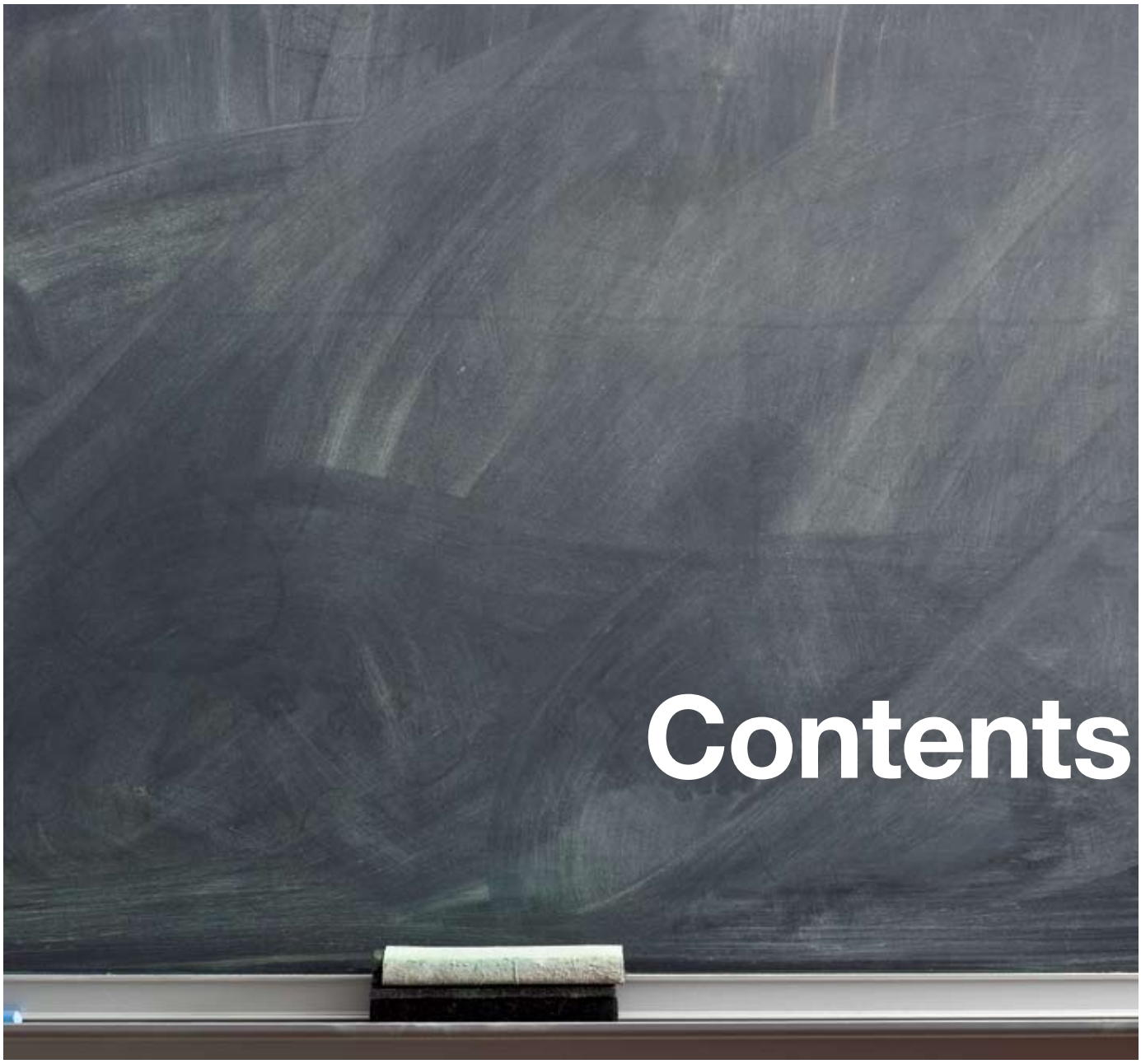
# How the world's most improved school systems keep getting better.

# Authors

Mona Mourshed | Chinezi Chijioke | Michael Barber

The authors deeply thank the over 200 system leaders, staff, and educators whom we interviewed across the 20 systems during this research. We further acknowledge the following leaders and experts for their counsel and thought partnership: KK Chan, John Deasy, Michael Fullan, S. Gopinathan, Peter Hill, Alan Kantrow, Lee Sing Kong, Tom Payzant, Andreas Schleicher, and Tan Ching Yee. The authors are deeply grateful to the substantial and committed contributions of our colleagues Eman Bataineh and Hisham Zarka, and our editor Ivan Hutnik, without which this report would not have been possible. The following colleagues provided valuable input and interview support throughout our work: Akshay Alladi, Byron Auguste, Tara Azimi, Alexander Busarov, Li-Kai Chen, Marcos Cruz, Sidnei Franco, Andrew Moffit, Michael Okrob, and Ramya Venkataraman. Lastly, we thank Nicholas Dehaney for his design creativity.

# Acknowledgements



# Contents

Foreword 10
Preface 12
Introduction and Overview 14
The approach 17
Lots of energy, little light 20
How to get there from here 24

1. Intervention 30
Through the looking glass 33
It's a system thing, not a single thing 37
Prescribe adequacy, unleash greatness 52
Common but different 61

2. Contextualizing 68
Break through, rather than break down 71
The guiding principles in mandating versus persuading 71

3. Sustaining 80
Collaborative practice: The user interface 84
The mediating layer: The operating program 91
Architecting tomorrow: The CPU 97

4. Ignition 100
Getting going 103
Never waste a good crisis 105
Nowhere to hide 106
Entering stage right 109
The new leader's playbook 110
Staying power 115

5. Conclusion 120

Appendix 126
Footnotes 136

# Foreword

There is a recent and rapidly growing appetite for figuring out and accomplishing what I call "whole system reform"---how to improve all schools in a district, a region, a state, province of country. For a long time, there has been the realization that better education is the key to societal and global productivity and personal and social well-being. Only recently are we beginning to see that interest turn into specific questions about how you actually go about whole system reform. What pathways, from what starting points, are going to get results in reasonably short time frames? How do we actually 'raise the bar and close the gap' for all students?

*How the World's Most Improved School Systems Keep Getting Better*—a report that examines 20 systems in action-- makes a unique contribution to this critical global agenda. Building on their 2007 study but with much more precision, in this remarkable report McKinsey gets inside the pathways. It sorts out systems according to starting points and progression. These performance stage continua—from poor to fair, fair to good, good to great, and great to excellence—are in turn unraveled according to intervention clusters within given contexts. In each case it is very clear that all improving entities, even if their starting point is dismal, are led by a combinations of leaders who are self-aware that they are engaged in a phenomenon that the report calls 'it's a system thing'—a small number of critical factors that go together to create the chemistry of widespread improvement.

We see the clusters of interventions, different for those starting from a weak base than those who have already had significant success. We see the pathways playing themselves out in each type of context. We see what it takes to ignite system change, what specific strategies achieve breakthrough, what interventions build ever -increasing momentum, how systems can sustain improvement, and especially how they can go to the next stage of development.

As someone who has worked explicitly on system change in several contexts since 1997, including being directly involved in helping to lead whole system reform in Ontario since 2003, I can say that *How the World's Most Improved School Systems Keep Getting Better* makes a one of a kind seminal contribution to this dynamic and critical field. It couldn't come at a more propitious time. Finally, we are witnessing across the globe a robust anticipatory and proactive interest in OECD's Programme for International Student Achievement (PISA). PISA is no longer just a 'results phenomenon'. PISA leaders are increasingly getting at what lies behind the numbers and are thus generating key insights and questions. The *How the World's Most Improved School Systems Keep Getting Better* report goes further, much further, in portraying the inner workings of successful pathways of reform given different beginning points.

We don't have a perfect storm yet but there is one brewing. This report is invaluable for policy makers and school system leaders who are or should be crafting a roadmap for improving their specific systems. It furnishes a powerful analytical tool with its intervention data-base to help guide such action. It will stimulate a wave of further whole system reform efforts, and will be accompanied by an associated body of research that will help us assess and learn with very specific lenses provided by this report.

The world needs to become much more wise about what lessons to extract for systems at different starting points, both with regards to the 'what' and 'how' of system reform. This is no ordinary report. It has captured action in real time. It will, by its clarity and compelling insights, catapult the field of whole system reform forward in dramatic ways.

Michael Fullan
Professor Emeritus, University of Toronto
Special Advisor on Education to the Premier of Ontario

In 2007, McKinsey & Company wrote a report on the common attributes of excellent school systems titled, *How the World's Best-Performing School Systems Come Out on Top*. As we discussed its contents with policymakers and education leaders around the world, one question came up time and again: "How does a system with modest performance become great?" The leaders we spoke to also wanted to know which aspects of a school system reform journey are universal and which are context-specific. Bearing these questions in mind, we decided to dedicate another major research effort to understanding the transformation of school system performance around the world.

This report is the result of that effort.

Our focus here is in analyzing the experiences of 20 school systems from all parts of the globe that have achieved significant, sustained, and widespread gains, as measured by national and international standards of assessment. The Appendix describes our system selection criteria, as well as our database structure for the detailed evidence we gathered to map the experiences of nearly 575 reform interventions made across the school systems in our research sample. Our purpose in this work has been to understand precisely which interventions occurred in each school system and when, and how these interventions interacted with each other and with the system's broader context to deliver better outcomes for students.

In our sample we included school systems that have undertaken a journey of improvement along all the different stages of the performance spectrum – from poor to fair, from fair to good, from good to great, and from great to excellent[1]. This spectrum rests, in turn, on a universal scale of calibration that we developed by normalizing several different international assessment scales of student outcomes discussed in the education literature. Our findings

are not, however, the result of an abstract, statistical exercise. In addition to assessment and other quantitative data, they are based on interviews with more than 200 system leaders and their staff, supplemented by visits to view all 20 systems in action.

Along the way, we have had the great pleasure and honor of meeting with hard-working and talented system leaders and educators around the world, all of whom have generously given of their time and provided us with unvarnished insight into what it is that has improved their system. We have had many memorable moments during our field research – certain systems, with long improvement journeys, arranged for us to meet the architects of reform who led the school system during the past 15-25 years (often pulling them out of retirement to do so). In other systems, ministers of education and heads of teacher unions came together in the same room to provide us with a full and transparent view of the collaborations and tensions in their improvement journey; in yet other systems districts and schools were opened to us so that we could hear directly the perspectives from the front line. Many system leaders used vivid language to describe the journey their school system had undergone: in Lithuania we heard of the "soup," while in Hong Kong we were told of the "typhoon." We thank all the people we have met during the course of this research and hope that we have accurately reflected their many insights.

We have taken the approach we have in this report in order to be able to support policymakers, school system leaders, and educators in understanding how systems with starting conditions similar to their own have charted a path to sustained improvement. In sharing the lessons of such experience, we hope that the children of the world will be the ultimate beneficiaries of their collective effort in crafting school improvement.

# Introduction and Overview

Almost every country has undertaken some form of school system reform during the past two decades, but very few have succeeded in improving their systems from poor to fair to good to great to excellent. This report looks closely at 20 school systems from different parts of the world, and from an array of starting points, that have registered significant, sustained, and widespread student outcome gains, and examines why what they have done has succeeded where so many others failed. In undertaking this research, we have sought to understand which elements are specific to the individual system and which are of broader or universal relevance. We believe that what we have discovered will help other systems and educational leaders to replicate this success.

There are few things as important to the future well-being of our world than the quality of the education our children receive. This is an important motivator for the vast majority of the leaders of the world's school systems. In speaking with these architects of school system reform, it became clear to us that not only are these leaders highly motivated and dedicated in genuinely trying to improve student outcomes but that they are hungry for more information on how to do so more effectively. We hope this report will go some way to meeting this desire.

It has been our assumption from the outset that the world's educational system reformers undertake improvement interventions that seem entirely plausible given their system context. During our interviews, the leaders of improving school systems all agreed that creating improvement required discipline and constant forward momentum. However, even amongst this august group, few were certain about why they had been successful: they often did not have a "theory of the case" about why what they did worked. Even fewer had a mental map of how all the changes they made fit together as a coherent whole. Some even thought they had just been lucky.

The lack of an overview is not surprising: education systems are inherently very complex and necessarily address disparate goals. Because no two systems face exactly the same challenges, it is very difficult to draw parallels between them or to see the wood for the trees. To add to this, school systems are constantly changing, so what worked a few years ago might well have little relevance today.

What our analysis reveals is that despite their different contexts, all improving school systems appear to adopt a similar set of interventions, one that is appropriate to their stage of the journey. This to not to say that context is not important, but it is secondary to getting the fundamentals right. This report attempts to disaggregate the various elements of what makes a school systems improve, to parse exactly what one system can learn from another, and how to adjust these elements to the specific, local context.

## The Approach

We followed a two-step process to select the school systems that form the subject of this research. First, we identified systems that have achieved significant, sustained, and widespread gains in student outcomes on international and national assessments from 1980 onwards. We differentiated these systems according to two categories, to ensure representation from both developed and developing country contexts. The first set, "sustained improvers," comprises systems that have seen five years or more of consistent rises in student performance spanning multiple data points and subjects; this group includes the systems of Singapore, Ontario, and Poland. The second set, "promising starts," are systems in developing countries or emerging areas that have begun data-supported reform efforts only recently, but which have already seen significant improvement over two to three years. The promising starts include the systems of Madhya Pradesh (India), Minas Gerais (Brazil), and Western Cape (South Africa). While the "promising starts" do not reach high attainment levels and few submit to international assessment, they have embarked on large-scale reform journeys employing innovative techniques that have shown significant (and sometimes remarkable) improvements in national assessments within a short period of time.

The second step was to select a broad and diverse set of systems from this improving group (Exhibit 1). Our sample comprises systems both large and small, centralized and decentralized, public and private. They are found on five continents and represent a wide array of starting performance levels (Exhibit 2). The Appendix describes our methodology in detail. ➔

# Exhibit 1: Our school system sample comprises 'sustained improvers' and 'promising starts'

**Sustained improvers**

Systems that have sustained improvement with 3 of more data points over 5 or more years

**Promising starts:**

Systems that have started improving as represented by ongoing improvement with just 2 data points or less than five years of improvement

| Systems | Time period of assessment | Sustained improvers | Promising starts |
|---|---|---|---|
| 1. Singapore | 1983 – 2007 | ✓ | |
| 2. Hong Kong | 1983 – 2007 | ✓ | |
| 3. South Korea | 1983 – 2007 | ✓ | |
| 4. Ontario, Canada | 2003 – 2009 | ✓ | |
| 5. Saxony, Germany | 2000 – 2006 | ✓ | |
| 6. England | 1995 – 2007 | ✓ | |
| 7. Latvia | 1995 – 2007 | ✓ | |
| 8. Lithuania | 1995 – 2007 | ✓ | |
| 9. Slovenia | 1995 – 2007 | ✓ | |
| 10. Poland | 2000 – 2006 | ✓ | |
| 11. Aspire Public Schools, USA | 2002 – 2008 | ✓ | |
| 12. Long Beach, CA, USA | 2002 – 2009 | ✓ | |
| 13. Boston/Massachusetts, USA[1] | 2003 – 2009 | ✓ | |
| 14. Armenia | 2003 – 2007 | | ✓ |
| 15. Western Cape, South Africa | 2003 – 2007 | | ✓ |
| 16. Chile | 2001 – 2006 | | ✓ |
| 17. Minas Gerais, Brazil | 2003 – 2008 | | ✓ |
| 18. Madhya Pradesh, India | 2006 – 2008 | | ✓ |
| 19. Ghana | 2003 – 2007 | | ✓ |
| 20. Jordan | 1999 – 2007 | | ✓ |

1 Primary focus was on Boston, within the context on Massachusetts State Reforms. Massachusetts NAEP results also indicate the state as a sustained improver from 1998-2007 on mathematics and reading

Source: McKinsey & Company interventions database

# Exhibit 2: Our selected systems represent a diverse mix

SOURCE: McKinsey & Company interventions database

At the heart of our analysis is a very extensive database. We asked improving systems to chronicle all the main interventions they undertook during the reform time period[2]; this ultimately yielded a database of almost 575 interventions across the 20 systems. We further categorized these interventions into ten areas of impact (e.g. professional development, accountability, learning model) and then disaggregated each of these ten areas into a total of 60 unique subareas: for example, the area of "accountability" includes the subareas of performance assessment, school inspections, and self-evaluation. We also categorized each intervention as to whether it constituted a change in structure, resource, or process, and in terms of which agent the intervention acted upon (e.g. principal, teacher, student).

In order to analyze the data, we first needed to be able to compare like with like. Collectively, the systems in our selection participated in 25 various international and national assessments[3] across multiple subjects (e.g. math, science, reading), school levels (e.g. primary and secondary), on a series of occasions, predominantly during the period from 1995 to 2010. Each of these assessments used a unique and independent scale. One of the critical underpinnings of this research has been to produce data that is comparable across the different systems over time and across assessments. To achieve this we used the methodology of Hanushek et al.[4] to normalize the different assessment scales on a single universal scale. Once the data had been normalized, we were able to classify the school systems' performance levels into four broad groupings across time: poor, fair, good, great, or excellent. We then mapped each system, with its interventions, onto a performance stage (poor to fair, fair to good, good to great, and great to excellent) and analyzed the intervention patterns revealed by the data.

What follows is a summary of the broad findings arising from this analysis. These findings are discussed in more detail in the following chapters of this report.

## Lots of energy, little light

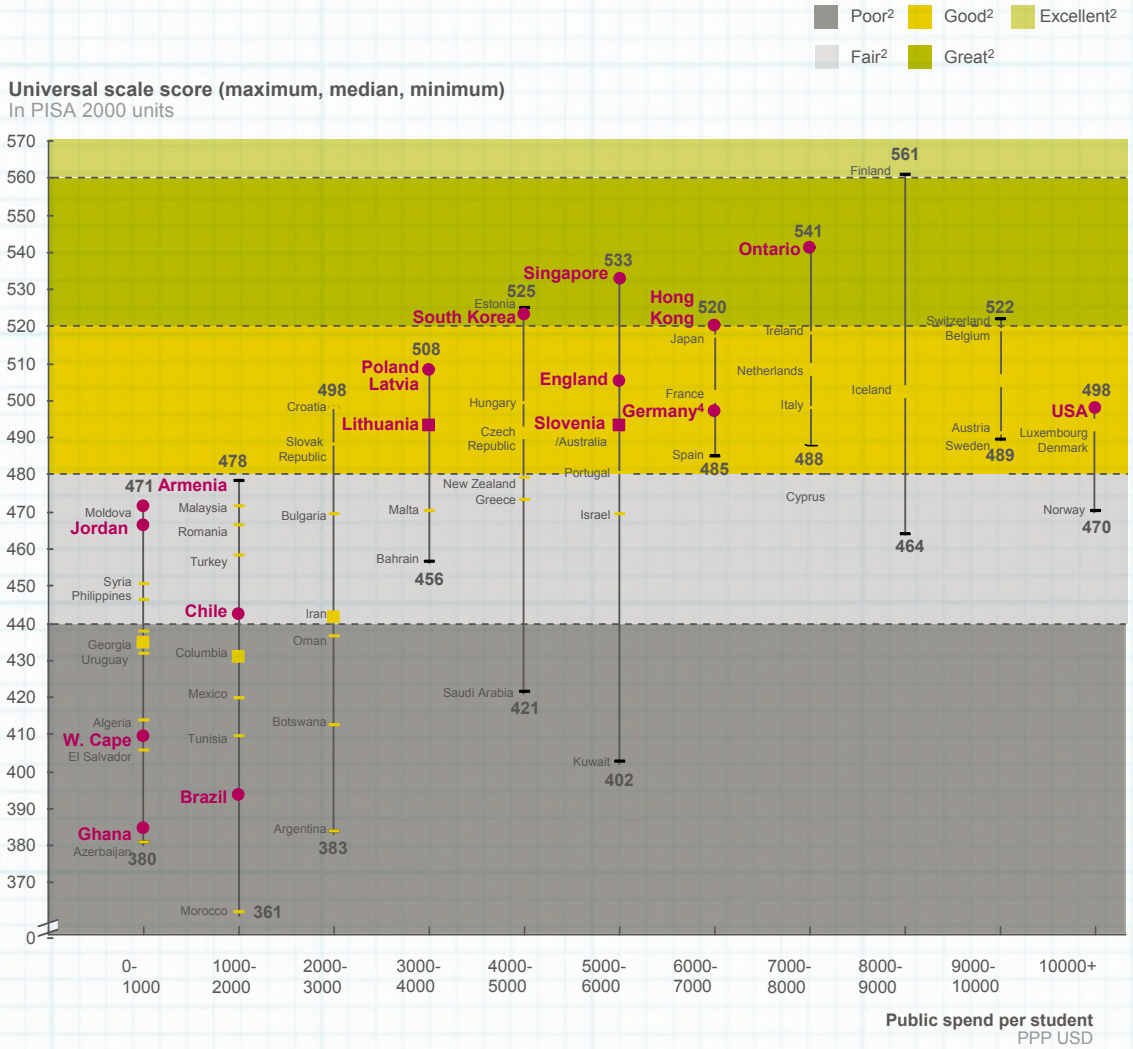
As we noted in our earlier report *How the World's Best-Performing School Systems Have Come Out on Top*, most OECD countries doubled and even tripled their spending on education in real terms between 1970 and 1994[5]. Unfortunately, despite this increase in expenditure, student outcomes in a large number of systems either stagnated or regressed. Moreover, based on the universal scale data, we find that systems with similar education spending have widely varying levels of performance – until the USD 6,000 spend per student (PPP) mark is reached, system performance spans the full spectrum of poor, fair, good, and great (Exhibit 3).

A few rays of hope penetrate this bleak landscape: in contrast to the majority, the school systems selected for our research sample have consistently improved student performance, as measured by national and international assessments, showing a steady upward trajectory for student outcomes over a period of ten years or more (Exhibit 4). Our sample systems are distinguished from other systems in that they achieve more with similar (or fewer) resources.

The systems focused on in this research demonstrate that significant improvement in educational attainment can be achieved within as little as six years (Exhibit 5). Their success does not simply attach to factors of wealth, scale, or even political system. Their improvements have been achieved irrespective of the individual system's starting point. For example, Hong Kong had a GDP per capita (at PPP) of over USD 42,000 and Latvia of USD 18,000. Saxony has 1,480 schools and Chile has 11,800 schools.

The lack of sustained progress seen in most school systems despite their massive investments should not be seen as the justification for abandoning the desire for educational improvement, but we believe it does demonstrate the need for adopting a different approach – one that will hopefully be guided by the experiences of school systems that have succeeded in improving over the longer term. ➔

# Exhibit 3:
# Systems with similar spend have widely ranging levels of performance



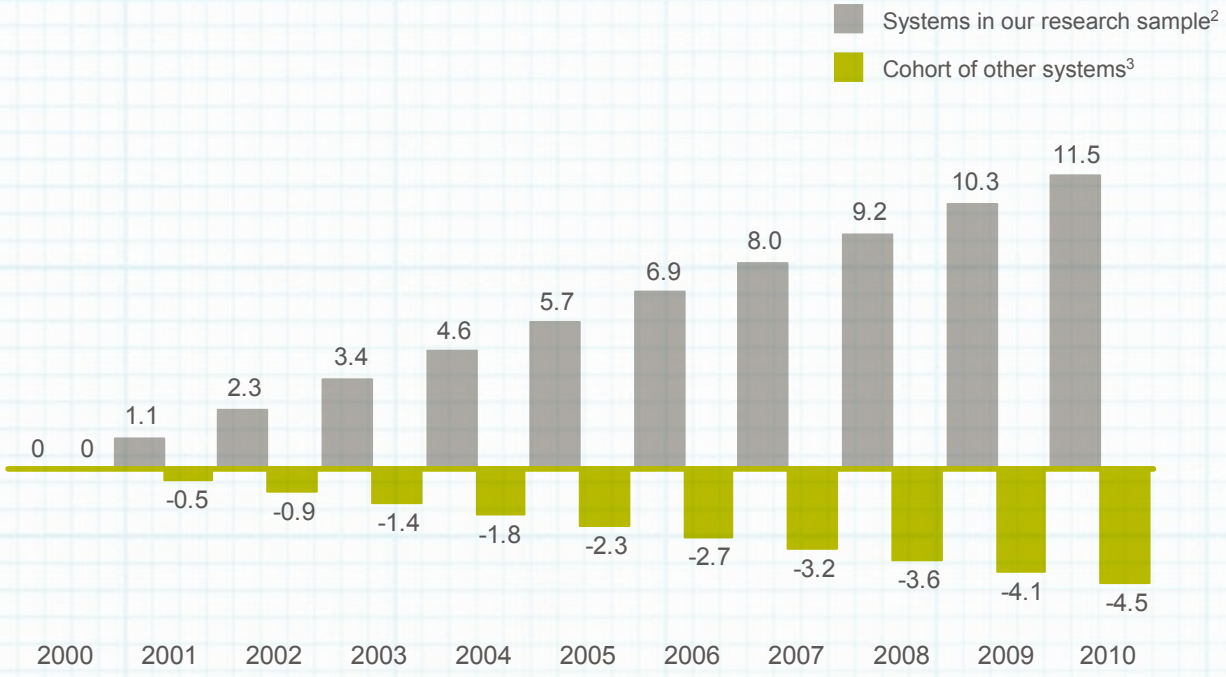

1 Universal scale created by McKinsey & Company, based on Hanushek R. Woessman methodology, to enable comparison across systems
2 Score cutoffs: Excellent >560; Great 520-560; Good 480-520; Fair 440-480; Poor <440
3 India (Madhya Pradesh) excluded due to lack of international assessment data
4 Saxony's universal scale score of 535 makes it "great" performing. Saxony is the highest performing state in Germany on 2006 PISA-E

Source: World Bank EdStats; IMF; UNESCO; PISA, TIMSS, PIRLS, McKinsey & Company

# Exhibit 4:
# Most school systems have stagnated or regressed in achievement, while our research sample has shown a steady upwards trajectory

**Trend of scores on universal scale since 2000[1]**

Increase since 2000 in units of PISA 2000




1 Trend line is the regression of average scores on the universal scale; SE of slope for the selected sample and the cohort are 0.40 and 0.43 respectively.

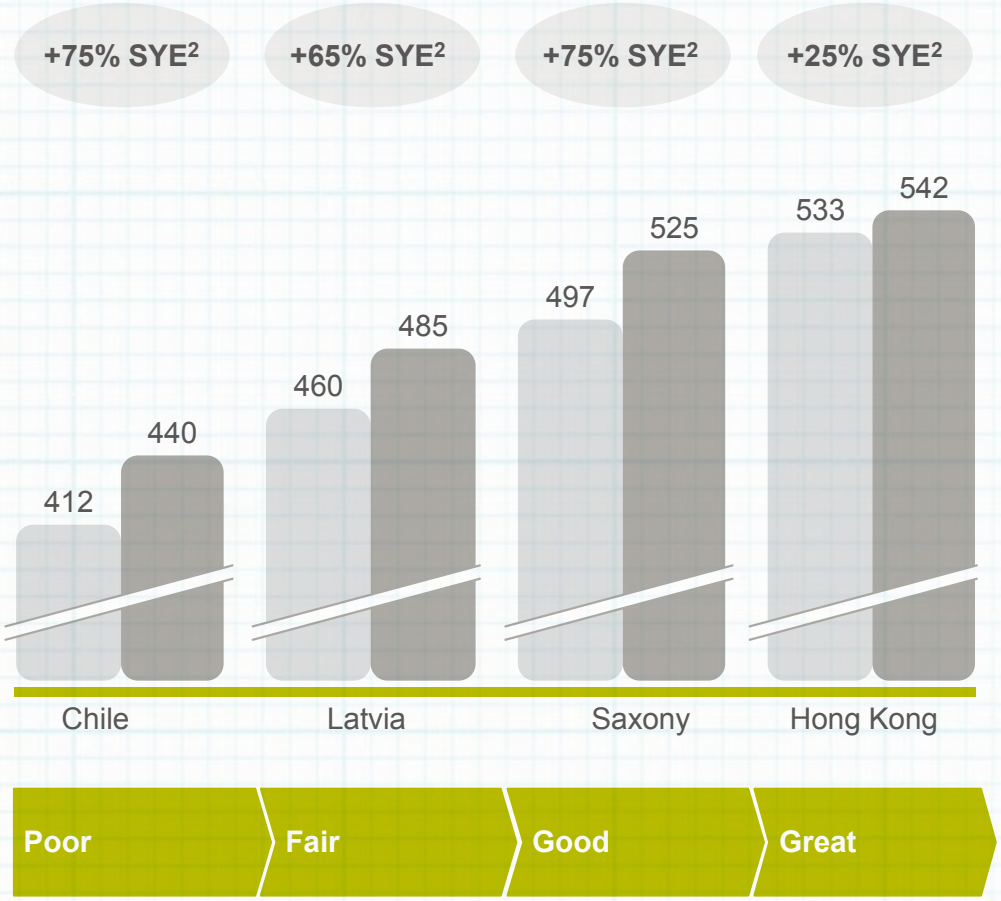
2 'Sustained improver' systems from our research sample, including: England, Hong Kong, Korea, Latvia, Lithuania, Ontario, Poland, Saxony, Singapore, Slovenia, Boston, Long Beach. Excludes 'promising start' systems and those for which special assumptions were made for the universal scale (e.g. Aspire)

3 A stable cohort of 43 systems not in our sample but participating in at least 3 assessments since 1999 were chosen, comprising: Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chinese Taipei, Colombia, Cyprus, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Indonesia, Iran, Ireland, Israel, Italy, Japan, Liechtenstein, Luxembourg, Malaysia, Mexico, Morocco, Netherlands, New Zealand, Norway, Portugal, Romania, Russia, Scotland, Slovak Republic, Spain, Sweden, Switzerland, Thailand, Tunisia, Turkey, USA

SOURCE: TIMSS, PISA, NAEP, national and provincial assessments, McKinsey & Company interventions database

# Exhibit 5:
# Systems at all performance levels can improve outcomes substantially in as short as six years

PISA scores, average[1]; 2000–2006




1 Average across math, science, and reading PISA scores
2 One school-year-equivalent (SYE) corresponds to 38 points on the PISA scale

Source: PISA, McKinsey & Company interventions database

## How to get there from here

What has confused much of the discussion about system improvement in the past is that each system's journey is different: each school system starts from a different point, faces different expectations, and operates in a different social and political context. These differences have often led even the experts to give poor advice. Rather like in the hoary old tale of a weather-worn farmer who, when asked directions by a lost traveler, replies, "Well I wouldn't be starting from here, if I were you." School system leaders, when looking for direction, are all too often told what to do from a starting point that is different from their own. Educators in a moderately performing system would be better off in seeking inspiration from similar systems that are managing to improve, rather than from those that are configured and positioned very differently, even if they are the world's best-performing ones.

This report shows that a school system can improve from any starting point. Its main message is that in order to do so, system leaders must integrate three aspects when developing and implementing an improvement journey. The first aspect is the status quo, called here the performance stage, which identifies the point where the system currently stands according to student outcomes. The second is the set of interventions necessary to make the desired improvements in student outcomes, here called the intervention cluster. The third is the system's adaptation of the intervention cluster to the prevailing context: taking into account the history, culture, politics, and structure of the school system and the nation.

We find that each performance stage is associated with a dominant cluster of interventions, irrespective of geography, culture, or political system. This comprises the set of interventions that systems use to successfully traverse from one stage to the next (e.g. from poor to fair). While the context does influence the emphasis and combination of interventions the system chooses from within this cluster, the intervention pattern is strikingly consistent for systems pursuing similar outcomes. However, we also find great variation in how a system implemented the same interventions, be it in terms of the sequence, the emphasis, or the rollout approach across schools. It is in contextualizing the intervention cluster where we saw the impact of history, culture, structure, and politics come fully into play.

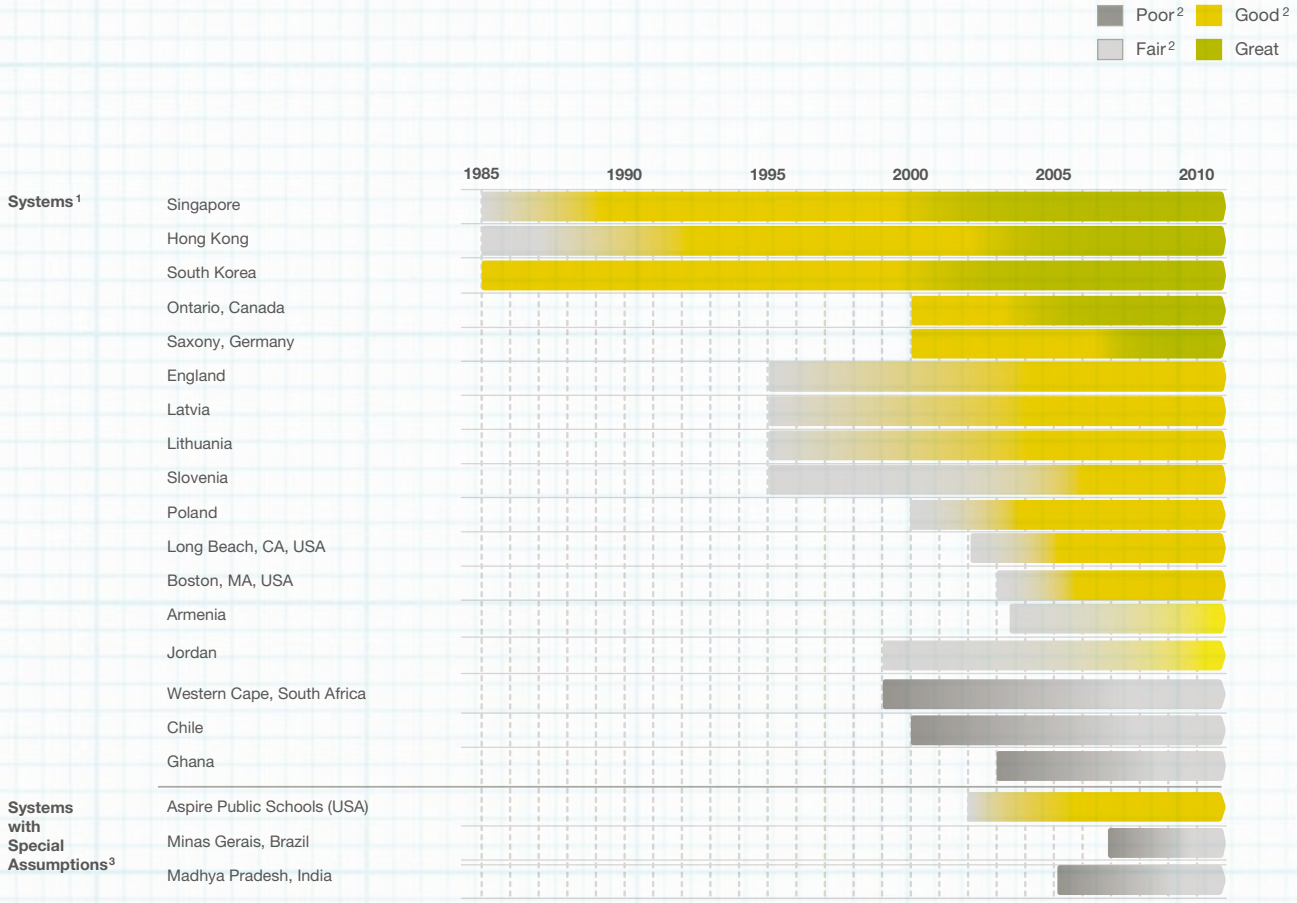
To complete our picture of the complex landscape of school system improvement journeys, in addition to the three basic elements – performance stage, intervention cluster, and contextualizing – we have added two more elements: sustaining and ignition. Sustaining is all about how a system puts in place the processes for ensuring improvement is continued over the longer term, and compromises three elements: the formation of a mediating layer between schools and the 'center', a strong pedagogy supported by collaborative practices; and leadership continuity. Ignition describes the conditions necessary to spur a system to embark on its reform journey. These conditions show remarkable consistency across all the improving systems studied here.

It needs to be kept in mind, that in the real world, each of these elements is integrated into a whole – the school system – just as human body or a car does not function as a collection of bits. Having acknowledged this, we will now focus on each of the bits, for it is in understanding their role that the functioning of the whole becomes clear.

### Performance stage

We have divided our 20 school systems that have been successful in sustaining improvement into performance stages. There are two important aspects to these stages. First, they are stages in two metaphorical senses of the word: reflecting how far the system has progressed relative to others; and the place or ground on which the interventions are acted out. Second, the performance stage is really a snapshot of a moment in time in a dynamic process. In actuality, each successful school system is undergoing a continuous progression from one performance stage to the next – an improvement journey. Exhibit 6 illustrates where our sample systems lie on the improvement continuum of poor to fair, fair to good, good to great, and great to excellent. As can be seen in the exhibit, some of the systems have moved all the way from fair to great, though over a period of many years. ➔

# Exhibit 6:
# Our sample represents a continuum of improvement from poor to fair to good to great




1 Systems were categorized across time as poor, fair, good, or great based on their average performance across test instrument, subject, and age group in a particular year on the universal scale. Systems also improve within each phase (e.g. England improved significantly within good but without crossing the threshold to great)
2 Score cut offs: Excellent >560 (none of our sample systems achieved this level); Great 520-560; Good 480-520; Fair 440-480; Poor <440.
3 No directly comparable assessment data to link these systems to international assessments exists, so special assumptions were made in placing them on scale.

Source: TIMSS, PISA, NAEP, national and provincial assessments; McKinsey & Company interventions database

## Intervention cluster

The school systems that have been successful in improving select an integrated set of actions from the menu of the interventions appropriate to their level of performance (see below). These improving systems appear to be careful in maintaining the integrity of the interventions; the evidence suggests that during each performance stage they select a critical mass of interventions from the appropriate menu and then implement them with fidelity. This is akin to the discipline of an exercise regimen – for the participant to be successful they need to be consistent in all its aspects, including diet and exercise, and in practicing these aspects regularly. The systems that have been unsuccessful in trying to improve may carry out the same types of interventions that successful systems undertake – but there appears to be one critical difference, that they are not consistent, either in carrying out the critical mass of interventions appropriate to their performance stage, or in pursuing them with sufficient rigor and discipline.

We have identified two different types of interventions carried out by improving school systems: the first set of interventions are those that are appropriate to a particular performance stage; the second set of interventions applies equally during all stages, but manifests differently in each stage.

1. Stage-dependent interventions: it's a system thing, not a single thing. These sets of interventions vary from stage to stage. Each set is discrete and is sustained throughout the stage.

- Poor to fair: the interventions in this stage focus on supporting students in achieving the literacy and math basics: this requires providing scaffolding for low-skill teachers, fulfilling all basic student needs, and bringing all the schools in the system up to a minimum quality threshold.

- Fair to good: at this stage the interventions focus on consolidating the system foundations; this includes the production of high quality performance data, ensuring teacher and school accountability, and creating appropriate financing, organization structure, and pedagogy models.

- Good to great: the interventions at this stage focus on ensuring teaching and school leadership is regarded as a full-fledged profession; this requires putting in place the necessary practices and career paths to ensure the profession is as clearly defined as those in medicine and law.

- Great to excellent: the interventions of this stage move the locus of improvement from the center to the schools themselves; the focus is on introducing peer-based learning through school-based and system-wide interaction, as well as supporting system-sponsored innovation and experimentation.

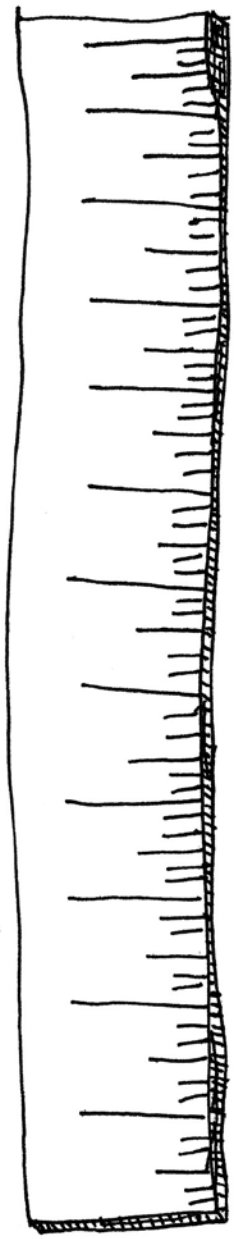
We further observe a correlation relationship between a system's performance stage and the tightness of central guidance to schools. Improving systems "prescribe adequacy but unleash greatness."[6] Systems on the journey from poor to fair, in general characterized by less skilled educators, tightly control teaching and learning processes from the center because minimizing variation across classrooms and schools is the core driver of performance improvement at this level. In contrast, the systems moving from good to great, characterized by more highly skilled educators, provide only loose guidelines on teaching and learning processes because peer-led creativity and innovation inside schools becomes the core driver for raising performance at this level.

2. Cross-stage interventions: common but different
   The cross-stage interventions comprise a group of six actions that occur with equal frequency across all performance stages, but manifest differently in each one. These six interventions are: revising the curriculum and standards, ensuring an appropriate reward and remunerations structure for teachers and principals, building the technical skills of teachers and principals, assessing students, establishing data systems, and facilitating improvement through the introduction of policy documents and education laws.

## Contextualizing

School systems that sustain improvement over the longer term have learned both how to navigate the challenges of their context and to use their context to their advantage. The leaders of these systems

tailor the three types of interventions required to their system's performance stage and circumstances. Contextualizing is all about the tactics the system leaders use in tailoring the set of the interventions needed on their performance journey to their specific context. Our research shows that the system leaders' prime aim in contextualizing the interventions is usually to gain the requisite support of the various stakeholders for the interventions being made.



In talking to leaders and architects of the improving systems, it appears that one of the biggest choices facing school systems when contextualizing their interventions is to what degree an intervention should be mandated and to what extent should persuasion be used. The systems we studied have adopted different combinations of mandating and persuading to implement the same set of interventions. These choices appear to be based on four contextual attributes: 1) the desired pace of change; 2) whether the desired change is a "non-negotiable" for the system reform; 3) the degree to which there are stark winners and losers as a result of the change; and 4) the credibility and stability of the system leadership and national government, and the historical and political context.

## Sustaining

The sustaining practices of the new pedagogy are characterized by the internalization of teaching practices. They are not merely about changing the explicit structure and approach of the system, but about how teachers think about teaching. In the words of Lee S. Shulman, professional pedagogues recognize "an implicit structure, a moral dimension that comprises a set of beliefs about professional attitudes, values, and dispositions."[7] We have found that there are three ways that improving systems commonly do this: by establishing collaborative practices between teachers within and across schools, by developing a mediating layer between the schools and the center, and by architecting tomorrow's leadership.

Many systems in our sample have created a pedagogy in which teachers and school leaders work together to embed routines that nurture instructional and leadership excellence.[8] They embed routines of instructional and ➔

leadership excellence in the teaching community, making classroom practice public, and develop teachers into coaches of their peers. These practices are supported by an infrastructure of professional career paths that not only enable teachers to chart their individual development course but also make them responsible for sharing their pedagogical skills throughout the system. In general, collaborative practices shift the drive for change away from the center to the front lines of schools, helping to make system improvement self-sustaining.

As the school systems we studied have progressed on their improvement journey, they seem to have increasingly come to rely upon a "mediating layer" that acts between the center and the schools. This mediating layer sustains improvement by providing three things of importance to the system: targeted hands-on support to schools, a buffer between the school and the center, and a channel to share and integrate improvements across schools. As our sample systems have moved through their improvement journey, a number have chosen either to delegate responsibility away from the center to a newly created mediating layer located between the central educational authority and the schools themselves (e.g. school clusters or subject-based groups), or have expanded the rights and responsibilities of an existing mediating layer (e.g. school districts/regions).

The third element commonly witnessed in sustaining school system improvement is the continuity of the system's leadership. This plays an important role in ensuring that the priorities, drive, mindset and resourcing of change is sustained across leaders. All systems need to somehow traverse smoothly from one leader to the next, so that change becomes evolutionary in nature. The most successful systems actively foster the development of the next generation of system leadership from within, ensuring that there is a continuity of purpose and vision in sustaining the system's pedagogy and improvement.

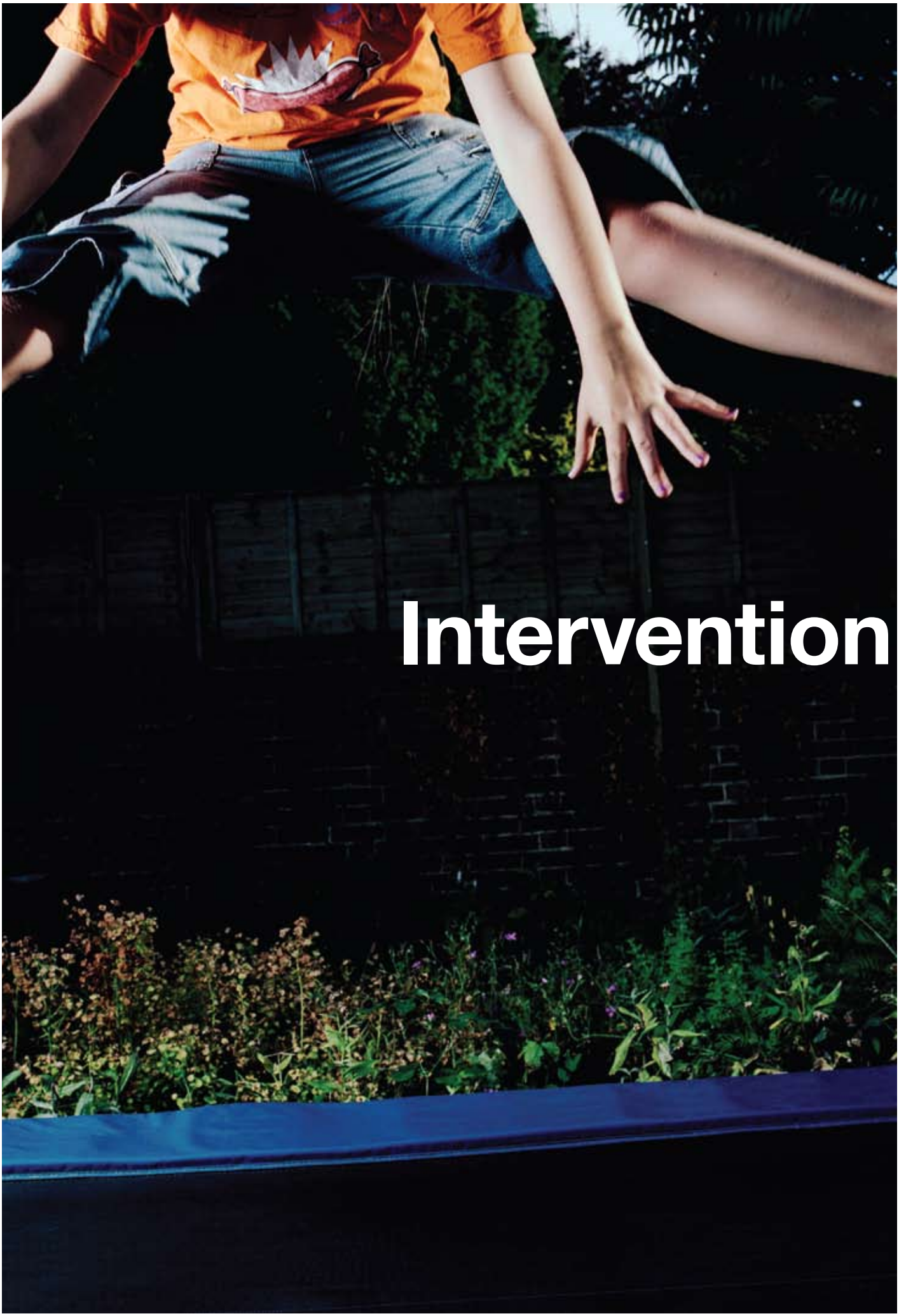
## Ignition

The question many might well ask at this point is, "How do we get started?" The starting point for every system embarking on an improvement journey is to decide just how to overcome the present inertia. Across our sample systems, the impetus required to start school system reforms – what we call ignition – resulted from one of three things: the outcome of a political or economic crisis, the impact of a high-profile, critical report on the system's performance, or the energy and input of a new political or strategic leader.[9] We find that fifteen out of our 20 studied systems had two of these ignition events present prior to the launch of their reform efforts.

Of the three, however, the injection of new leadership appears to be by far the most important factor: all 20 of the systems studied here have relied upon the presence and energy of a new leader to jumpstart their reform program. New technical leaders were present in all of our sample systems, and new political leaders present in half. These new leaders tend to follow a common "playbook" of practices upon entering office. Once installed, they have staying power: the median tenure of the new strategic leaders is six years and that of the new political leaders is seven years, thereby enabling continuity in the reform process and development of the system pedagogy. This is in stark contrast to the norm. For example, the average tenure for superintendents of urban school districts in the U.S. is nearly three years; the tenure of education secretaries in England is just two years on average, similar to that of education ministers in France.

It is clear from what we have said here that while there is no single path to improving school system performance, the experiences of all 20 improving school systems – both the "sustained improvers" and the "promising starts" – have strong commonalities in the nature of their journeys. We hope this analysis will provide system leaders with the opportunity to rigorously assess where their system is on its path to improvement and to what extent they are already making use of the appropriate set of interventions – and whether there might be the opportunity to do things differently.

What follows in the main part of the report explores each of the various dimensions of the school system performance journey in more detail. The report is divided into four chapters: interventions, contextualizing, sustaining, and ignition. ■



# Intervention

One of the major challenges for each and every school system is to decide what interventions it should make in order to improve its performance. Based on our database of nearly 575 interventions mapped over time across our 20 sample systems, we can make three observations about the pattern of interventions undertaken:

1. We observe dominant clusters of interventions that all improving systems carry out at each journey stage on the long path from poor excellent;
2. There is a correlation between a system's performance level and the degree of tightness of central control over its school processes; and,
3. Six interventions occur with equal frequency across all journeys, but are manifested differently in each improvement journey stage.

## Through the looking glass

School system reform is a complex endeavor requiring system leaders to make decisions about numerous interlinked issues. In so doing, they have to take account not only of how to maintain their current system performance but also decide what interventions they will choose to make in order to improve that performance, while addressing the socio-economic, political, and cultural context within which they operate.

The question at the heart of our research is whether it is possible to produce a topographical route map for systems undertaking the journey required to transform their performance, one that will be useful in guiding them through this complexity. To this end, while our intent has been to fully embrace and appreciate the complexity of the decisions that improving system leaders need to make, we have focused on extracting a discernable pattern from their actions that could prove helpful to others.

Our analysis produced three main findings:

1. It's a system thing, not a single thing
   There is a common pattern in the interventions improving systems use to move from one performance stage to the next, irrespective of geography, time, or culture. These interventions, which we term the "improvement cluster," are mutually reinforcing and act together to produce an upward shift in the trajectory of the system. Though there is a different cluster of interventions for each stage of the system's journey (poor to fair, fair to good, good to great, great to excellent), there is a dominant pattern throughout that journey. ➔

## Methodology

**To understand whether there is a common intervention pattern or not, we asked the leaders of improving systems to chronicle all the main interventions they undertook in their systems over the reform period[10]. This ultimately yielded a database of almost 575 interventions across the 20 systems (what we refer to as the "interventions database"). We then categorized these interventions into ten broad areas (e.g. professional development, accountability, learning model), before further disaggregating each of the ten into a total of 60 distinct subareas. Accountability, for example, is an area including the subareas of performance assessment, school inspections, and self-evaluation. We then categorized each intervention according to whether it constituted a change in structure, resource, or process, and which agent (e.g. principal, teacher, student) it acted upon. We developed a universal scale of student outcomes to plot all systems onto a single achievement scale across time[11]. We mapped each system, with its interventions against the various stages of the improvement journey (from poor to fair, fair to good, good to great, and great to excellent) and undertook a series of analyses regarding the intervention pattern.**

**In order to determine the cluster of interventions per improvement journey, we followed a three-step process. First, we calculated how often each of the 60 unique intervention subareas occurred in a given improvement journey. Second, we analyzed the relative importance of each intervention occurrence in that given improvement journey stage relative to the other improvement journeys. We then assigned each intervention to the improvement journey in which it was most concentrated. For example,while enrolment (comprised of the subareas: fulfillment of basic needs, increasing school seats, and provision of textbooks) constitutes just eight percent of the total number of interventions made in the "poor to fair" improvement journey, it is almost ten times more concentrated in this stage than in the other improvement journey stages. As such, we assigned it to the "poor to fair" journey. Lastly, we triangulated the analysis results with what we heard from system leaders during interviews about the most important interventions they undertook during their improvement journey.**
**The Appendix contains a detailed explanation of our methodology.**

2. Prescribe adequacy, unleash greatness
   There is a strong, correlation between a school system's improvement journey stage and the tightness of central control over the individual schools activities and performance. Systems on the poor to fair journey, in general characterized by lower skill educators, exercise tight, central control over teaching and learning processes in order to minimize the degree of variation between individual classes and across schools. In contrast, systems moving from good to great, characterized by higher skill educators, provide only loose, central guidelines for teaching and learning processes, in order to encourage peer-led creativity and innovation inside schools, the core driver for raising performance at this stage.

3. Common but different
   Our findings indicate that six interventions occur with equal frequency across all the improvement journeys, though manifesting differently in each one. These six interventions are: revising curriculum and standards, ensuring an appropriate reward and remuneration structure for teachers and principals, building the technical skills of teachers and principals, assessing students, establishing data systems, and facilitating the improvement journey through the publication of policy documents and implementation of education laws.

## To what extent can a system leader exercise choice?

Ultimately, every system leader is faced with the challenge of integrating three dimensions of the system's improvements in order to successfully develop and implement its improvement journey: its current level of performance, the necessary interventions, and the context in which these are made (Exhibit 7). The important question is to what extent a system leader can exercise choice in this algorithm? A simple answer is that all the improving systems we examined within a given journey show little variation in what they do, but a much greater extent of variation in *how* they do it.

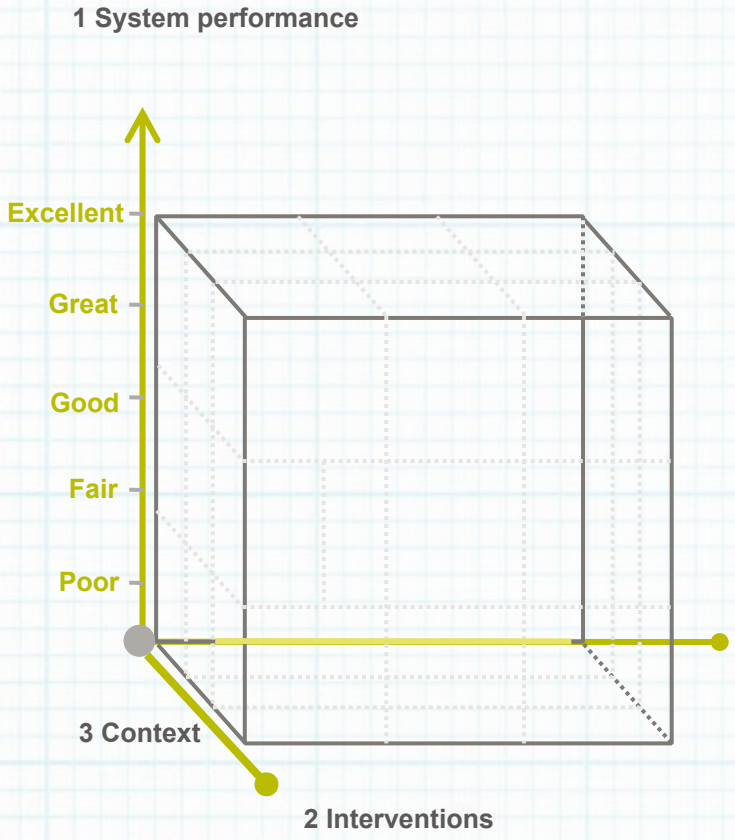
The evidence suggests that each journey stage comes equipped with a dominant intervention cluster – this is the sum total of individual interventions we observed systems using to raise the level of their performance from one stage to the next. The intervention cluster can be thought of as a menu from which the improving systems implement a critical mass.

This is not to suggest that systems have no choice: they have a great deal of choice in how they implement these interventions, in terms of the sequence, the emphasis, or the manner in which the system rolls out the interventions across its schools. It is here that we see the impact of history, culture, structure, and politics come fully into play, producing significant differences in the particulars of how systems manifest their reforms. Chapter 3 explores the contextualizing of interventions in depth.

To use a simple analogy, a person seeking to lose weight sustainably must do two things: exercise and consume fewer calories. They must do both for the regimen to be fully effective. These two interventions are akin to the intervention cluster, and are true irrespective of where this person lives in the world. Once the regimen has been embarked upon, this person now has the choice of how to implement the exercise program (tennis, hiking, gym, etc.) and diet (all protein, balanced blend of carbohydrates and protein, liquid, etc.). Their decision about which combination to follow will and should be based on their personal preferences, metabolic rate, and attributes; otherwise, they will quickly abandon their weight loss plan. This is where culture and tradition play a key role. Similarly, though there is a dominant cluster of interventions for each improvement journey stage, system leaders must then decide on an implementation path that suits its context in order to be able to sustain and persevere with its improvement program. Willpower, discipline, and persistence are required to see both weight loss and school system reform through to transformation.

Though there is no magic formula for improving school system performance, this research points to a clear path that improving systems have undertaken at each stage in their journey – a path illuminated by signposts. The remainder of this chapter describes this path and its signposts in greater detail. ➔

## Exhibit 7:
## A system leader must integrate three dimensions when crafting and implementing an improvement journey




**1 Assess current performance level**

- Measure student outcomes
- Decide if current level is poor, fair, good, great, or excellent

**2 Select interventions**

- Decide what the system needs to do in order to raise student outcomes, guided by its performance level and specific challenges

**3 Adapt to context**

- Tailor leadership style and tactics (e.g. mandate or persuade) to the history, culture, politics, structure etc. of the school system and nation

Source: McKinsey & Company

# Exhibit 8:
# A unique "intervention cluster" exists for each improvement journey, with six interventions common across all journeys

| Improvement journey | Poor to fair | Fair to good | Good to great | Great to excellent |
|---|---|---|---|---|
| Theme | *Achieving the basics of literacy and numeracy* | *Getting the foundations in place* | *Shaping the professional* | *Improving through peers and innovation* |
| Intervention cluster[1] | ▪ **Providing motivation and scaffolding for low skill teachers**<br>– Scripted teaching materials<br>– Coaching on curriculum<br>– Instructional time on task<br>– School visits by center<br>– Incentives for high performance<br>▪ **Getting all schools to a minimum quality level**<br>– Outcome targets<br>– Additional support for low performing schools<br>– School infrastructure improvement<br>– Provision of textbooks<br>▪ **Getting students in seats**<br>– Expand school seats<br>– Fulfil students' basic needs to raise attendance | ▪ **Data and accountability foundation**<br>– Transparency to schools and/or public on school performance<br>– School inspections and inspections institutions<br>▪ **Financial and organisational foundation**<br>– Optimization of school and teacher volumes<br>– Decentralizing financial and administrative rights<br>– Increasing funding<br>– Funding allocation model<br>– Organizational redesign<br>**Pedagogical foundation**<br>▪ – School model/streaming<br>– Language of instruction<br>– Language of instruction | ▪ **Raising calibre of entering teachers and principals**<br>– Recruiting programs<br>– Pre-service training<br>– Certification requirements<br>▪ **Raising calibre of existing teachers and principals**<br>– In-service training programs<br>– Coaching on practice<br>– Career tracks<br>– Teacher and community forums<br>▪ **School-based decision-making**<br>– Self-evaluation<br>– Independent and specialized schools | ▪ **Cultivating peer-led learning for teachers and principals**<br>– Collaborative practice<br>– Decentralizing pedagogical rights to schools & teachers<br>– Rotation and secondment programs<br>▪ **Creating additional support mechanisms for professionals**<br>▪ Release professionals from admin burden by providing additional administrative staff<br>▪ **System-sponsored experimentation/innovat ion across schools**<br>– Providing additional funding for innovation<br>– Sharing innovation from front-line to all schools |
| Common across all journeys | **Six interventions: [1] Revising curriculum and standards; [2] Reviewing reward and remunerations structure; [3] Building technical skills of teachers and principals, often through group or cascaded training; [4] Assessing student learning; [5] Utilizing student data to guide delivery, and [ 6] Establishing policy documents and education laws** | | | |

1 Total number of interventions in each phase: poor to fair, n=103, fair to good, n=226, good to great, n=150, great to excellent, n=94

Source: McKinsey & Company interventions database

## It's a system thing, not a single thing

As we examined the pattern of system interventions emerging from our research, we sought to test two hypotheses: 1) do the system's choices of interventions vary in the four improvement journey stages; and, 2) do the systems engaged in the same improvement journey exhibit the same, dominant intervention pattern, one that is consistent across geography, time, and culture.

Our interventions database supports both these hypotheses (Exhibit 8). The following sections describe the nature of these four improvement journey stages, and the intervention that characterize them.

### The "poor to fair" journey: achieving basic literacy and numeracy

The systems in our sample moving from poor to fair confronted five main challenges at the outset of their improvement journey. First, due to the challenges inherent in the place they start from, their teachers and principals were less experienced and less motivated than in systems further along the journey. Second, the governing education bodies had little capacity for supporting and managing schools; this problem was all the more acute due to the large size of many of these systems. Third, performance varied widely between schools in a particular system. Fourth, only limited resources were available for the improvement program (both human and financial). Fifth, the levels of student literacy and numeracy were low, and the level of absenteeism significant.

In addressing these challenges, we found that three of the systems, comprising Minas Gerais (Brazil), Madhya Pradesh (India), and Western Cape (South Africa), had sharply defined programs to raise basic literacy and numeracy outcomes, particularly at the primary level. Our field interviews further indicated that the leaders and stakeholders in these three systems could describe a well-defined path along which they were making progress. The other two systems in this journey, Chile and Ghana, also had the objective of raising literacy and numeracy, but by their leaders' own admission, were more focused on improving the system environment (e.g. ensuring adequate textbook provision, increasing student time given to the task) than in following a systematic program. Interestingly, while international assessments showed significant improvement for both Chile and Ghana, their system leaders were unclear about what exactly transpired in their system to result in this improvement.[12]

Despite the geographic and cultural diversity between the different systems in Madhya Pradesh, Western Cape, and Minas Gerais, all three selected a strikingly similar cluster of interventions in order to achieve their common goal to achieve rapid gains in basic literacy and numeracy outcomes at the primary level. Moreover, their intervention pattern and objectives mirrors those of systems that underwent their poor to fair journey in previous decades, such as that of Singapore during the 1970s and 1980s. Exhibit 9 describes the intervention cluster that they implemented.

The example of Minas Gerais, the third-largest state in Brazil, demonstrates how these interventions come together in holistic system improvement. In 2006, a state-wide assessment showed that only 49 percent of its eight-year-olds were able to read at the recommended level of proficiency. The governor set the aspiration that by 2010, 90 percent of eight-year-olds would read at the recommended level. This involved 2500 primary schools, 15,000 teachers and 500,000 students.

The state's department of education translated this overarching goal into specific regional and school-level improvement targets. A "results book," including baseline student achievement data, was created for each school so that teachers and principals could see their starting point and evaluate their progress. The Department of Education then developed prescriptive teaching materials for each lesson, to guide teachers in their classroom activities, and provided new workbooks for the students. The guides proved so effective that several private and municipal schools also voluntarily adopted the materials. It also strengthened its capacity across the 2,450 primary schools in the state, creating a central team of 46 members divided across the four regions. Each core team spent two weeks per month visiting ➔

# Exhibit 9:
# Poor to fair journeys focus on achieving basic literacy and numeracy

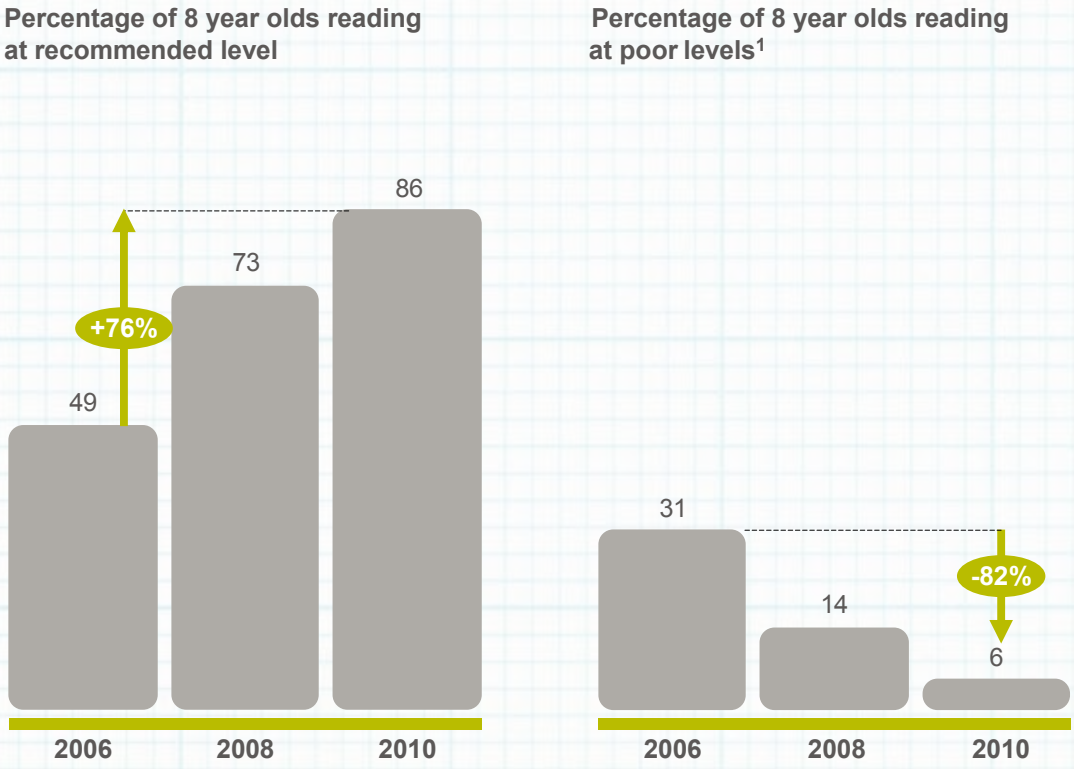
| Theme | Description | Example interventions |
|---|---|---|
| **Providing scaffolding and motivation for low skill teachers and principals** | ▪ **Scripted lessons:** The system creates instructional objectives, lesson plans, and learning materials for daily lessons to teachers lessons to enable teachers executing lessons rather than devising them<br>▪ **Coaching on curriculum:** The system creates a field force of coaches to visit schools and work with teachers in-class on effectively delivering the curriculum<br>▪ **Incentives for high performance:** The system gives rewards (monetary and prestige) to schools and teachers who achieve high improvement in student outcomes against targets<br>▪ **School visits by center**: The system's central leaders/administrators visit schools to observe, meet and motivate staff, and discuss performance<br>▪ **Instructional time on task**: The systems increases student instructional time | ▪ Prescriptive teaching materials<br>▪ Technical skill-building<br>▪ External coaches<br>▪ School visits by center<br>▪ Instructional time on task |
| **Getting all schools to minimum quality standard** | ▪ **Targets, data, and assessments:** The system sets minimum proficiency targets for schools/students, frequent student learning assessments (linked to lesson objectives, every 3-4 weeks), and data processes to monitor progress<br>▪ **Infrastructure:** The system improves school facilities and resources to a minimum threshold adequate for attendance and learning<br>▪ **Textbooks and learning resources:** The system provide textbooks and learning resources to every student<br>▪ **Supporting low performing schools:** The system funds targeted support for low performing schools | ▪ Outcome targets<br>▪ Assessments<br>▪ Data systems<br>▪ School infrastructure improvement<br>▪ Provision of textbooks<br>▪ Additional funding for low performing schools |
| **Getting students in seats** | ▪ **Expand seats:** The system increases school seats to achieve universal access<br>▪ **Fulfill students' basic needs:** The school provides for student basic needs to ensure that more students attend school and that absenteeism declines | ▪ Meeting basic needs (meals, clothing, transportation, toilets)<br>▪ Increase student seats |

**Systems included**

Chile (2001-2005)
Madhya Pradesh (2006+)
Minas Gerais (2003+)
Western Cape (2003+)
Ghana (2003+)

Source: McKinsey & Company interventions database and system interviews

# Exhibit 10:
# Following implementation of the literacy reform in 2006, Minas Gerais improved literacy levels and rose to the top of Brazil's national assessment




**From 2007 to 2009, Minas Gerais also rose from 5th place to 1st place among Brazilian states on Brazil's national (IDEB) assessments**

1 Poor performance level is defined by assessment as students are only able to read words

Source:: Brazil PROALFA reading assessment

regional departments of education assigned with three tasks: to train the trainers,[13] to disseminate and assess the implementation of the support materials developed by the Department of Education, and to act as a barometer and gather feedback from schools regarding their needs, challenges, and progress in implementing the literacy program.

Minas Gerais tracked the performance of each region, school, and student, putting in place an online database. The state-level core team used the analysis of this data to assess progress and differentiate its approach to its schools. It provided strong guidance and enforces tight accountability for schools with the largest target gaps, while allowing greater autonomy in higher performing schools, as long as they continued to meet targets. Teachers in schools that met their targets received up to one month's extra salary. Between 2006 and 2010, the percentage of eight-year-olds reading at the recommended level increased from 49 to 86 percent. During the same period, the number of students who were performing poorly dropped from 31 to 6 percent. By 2009 Minas Gerais had risen from fifth place to first in Brazil's National Education Index of student outcomes (Exhibit 10).

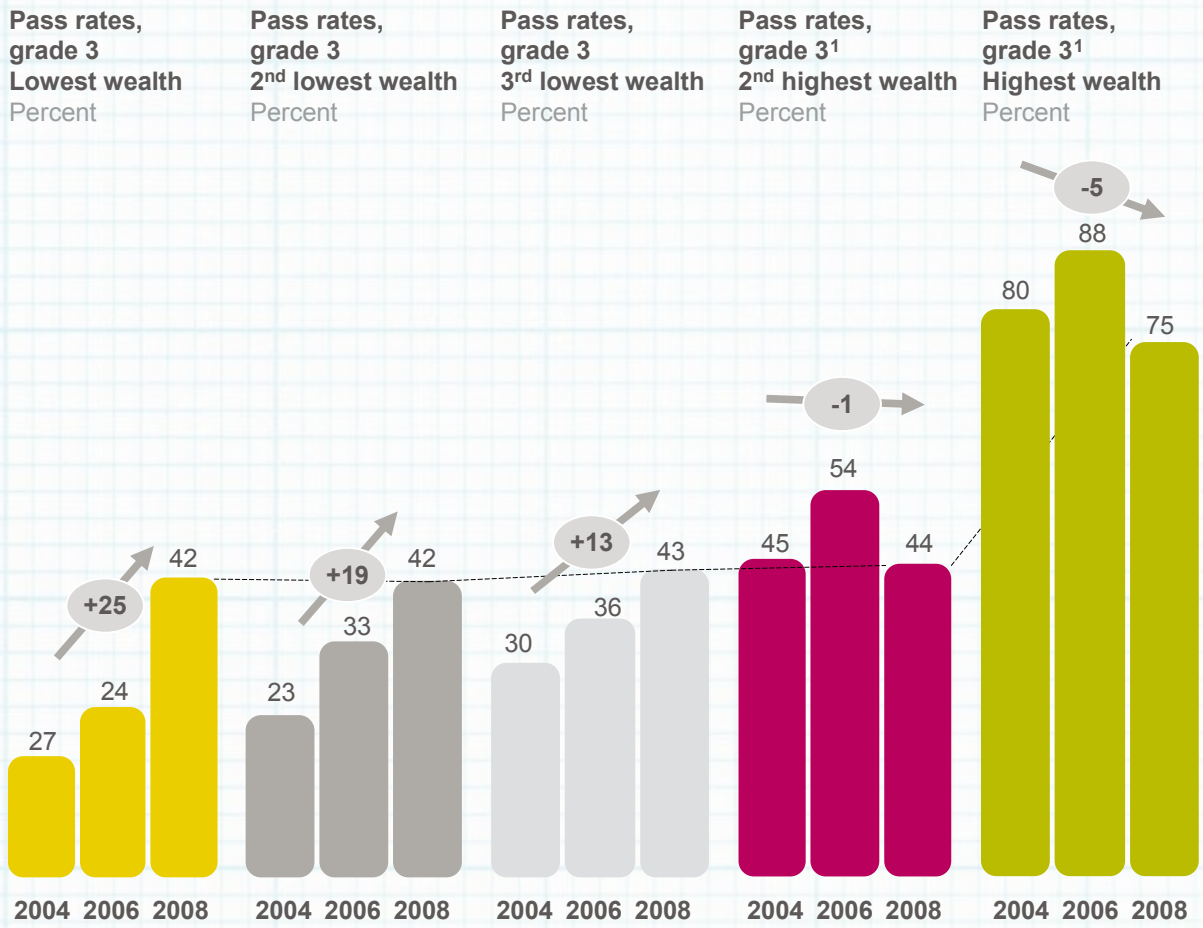
A critical achievement in the poor to fair improvement journey stage is to simultaneously raise overall outcomes while reducing performance variation across schools and socioeconomic groups. Western Cape (South Africa), for example, has achieved a steady rise in third and sixth-grade literacy levels since 2002, narrowing the achievement gap of the poorest and lowest performing quintiles of students. With regard to third-graders, the three quintiles from the lowest income group caught up with the second richest quintile over a period of four years (Exhibit 11). To achieve this improvement, the Western Cape Education Department (WCED) identified and developed strategies to support the lowest performers and raise the floor of outcomes. It combined data on school performance with geographic information in order to identify specific communities with performance challenges, understand the specific local needs of those communities and tailor its support accordingly. For example, in one district, district officers worked with illiterate parents to jointly write stories that they could memorize and recite to their children. It also asked the farm owners' association to allow farm workers (parents) time off to meet their children's teachers. WCED staffers spent three days annually with each of the eight districts in the state to review school performance data, speaking to the district leaders and parents, and visiting the highest and lowest-performing schools in the district.

Closing the achievement gap also commonly required two further interventions. First, the students' basic needs were met so that they could focus on learning. To this end, the Madhya Pradesh, Minas Gerais, and Western Cape programs all offered free school meals to their undernourished students. Additionally, Madhya Pradesh provided free uniforms and bicycles to improve enrolment and attendance, while some schools in Minas Gerais provided bathing facilities for their students. Second, the improving systems sought to increase the instruction time for literacy and numeracy. In Madhya Pradesh the timetable was altered so that two hours a day could be devoted to the new literacy lessons, for instance. Similarly, in Western Cape the system mandated 30 minutes a day for pleasure reading as part of its literacy improvement strategy.

The three systems' approaches were distinguished from each other by certain differences in style. In Madhya Pradesh a more regimented approach was taken in scripting and standardizing classroom teaching; interviewees attributed this to the enormity of the task; the state spans 138,500 public schools, 17 million students, and over 450,000 teachers. In contrast, Western Cape – with 1,100 primary schools, 600,000 students, and 17,000 teachers – allowed districts more flexibility in determining how they would get results. Aside from mandating 30 minutes a day to pleasure reading, the WCED did not stipulate any required instructional approach. However, in 2006, it tightened central guidance by requiring districts to address eight specific areas in their improvement strategy.[14]

Chile and Ghana, although having different contexts, focused more of their efforts on improving student attendance and in raising schooling standards to a minimum quality level. Ghana's ➔

# Exhibit 11:
# Western Cape narrowed the literacy inequality gap in four years: among 3rd graders, the bottom three quintiles have caught up to the second richest




1 Interviewees in WCED attributed some of the drop in the two highest wealth quintiles to shifts upwards in wealth categories of learners from 2006

Source: WCED Learner Assessment Studies, Final Reports, 2002-2008

main interventions included raising the coverage of primary education (net primary enrolment rose from 59 percent in 2004-05 to 89 percent in 2008-09); universal textbook distribution in core subjects (improving student-textbook ratio from 4:1 to 1:1); improving student health (providing de-worming, eye-screening, and potable water); the provision of free daily meals to deprived schools (to 20,000 schools in 2001-02, rising to 330,000 by 2007); and, in 2002, establishing nation-wide student assessments in order to provide schools with transparency on student performance. While some teacher capability-building occurred, it was not as systematic. In Chile, the flagship intervention was to expand the school day from one-half day to a full day in 1996, representing the equivalent of an additional two years in schooling for students. This additional time was used to teach content introduced in the recent curriculum reform; this aspect of the program was supported by efforts toward the universal provision of textbooks and learning materials, particularly in rural districts.

The evidence suggests that those systems on the poor to fair journey that were relentlessly focused on raising literacy and numeracy followed a common menu of interventions, whereas those systems focused on improving the overall system environment and structure (i.e. Ghana and Chile) were looser in their choice of interventions.

## The "fair to good" journey: consolidating the system foundations

Fourteen systems in our sample of 20 have journeyed from fair to good at some point in their recent history. Having achieved basic literacy and numeracy levels, these systems next sought to raise the quality of student skills. The critical issue they faced was how to configure the foundations of their system, including the creation of systems for data tracking, teacher accountability, finance, organization, and pedagogy. These foundations are essential for providing the systems with the necessary information, resources, and structures required to monitor and improve performance. Exhibit 12 describes the intervention cluster that characterizes the fair to good improvement journey.

Poland's experience illustrates the nature of this improvement journey. Prior to 1999, Poland had a school model comprising eight years of primary school and four years of secondary school; half of Poland's secondary students were placed on a vocational track and the other half on an academic track. The system leadership decided to increase general education by one year in order to provide a wider range of opportunity in secondary education. It therefore moved to a school model with six years of generalist primary education, three years of generalist lower secondary education, and three years of secondary school with academic, general, and vocational tracks. The structural and pedagogical implications of this decision were two-fold. First, Poland needed to create 4,000 lower secondary schools in one year, the vast majority of which were to be reconstituted from closing primary schools. The Ministry of Education tasked the municipalities with implementing this restructuring, allowing them to adopt approaches that were tailored to their local community context. Second, the Ministry created a new curriculum for lower secondary schools, which had implications for adjacent grades, and the need to train teachers accordingly.

In parallel, Poland decentralized the central government's administrative and financial power with regard to schools, as was consistent with Poland's overall decentralization drive. A strong belief existed across the system that the center could not effectively manage its schools from a distance. Poland therefore specified critical decision rights at each level of education – the center set standards, the regions (which the government consolidated from 49 to 16) inspected schools and provided pedagogical support; the districts controlled the administration and financing of secondary schools, while the municipalities controlled the administration and financing of primary and lower secondary schools. Lastly, at the school level, principals were able to choose which teachers to hire, while teachers could choose which curriculum to use from a pre-approved list of over a hundred private providers. Poland monitored the progress of the reform program by introducing national examinations at grades six, nine, and twelve, supplemented by annual students tests.

The other Eastern European and former Soviet states in our sample that are also engaged ➔

# Exhibit 12: Fair to good journeys emphasize getting the system foundations in place

| Theme | Description | Example interventions |
|---|---|---|
| **Data and accountability foundation** | ▪ **Transparency and accountability:** The system establishes student assessments and school inspections to create reliable data on performance and to hold schools accountable for improvement<br>▪ **Improvement areas:** The system uses this data to identify and tackle specific areas (e.g., subjects, grades, gender) with lagging performance | ▪ Student assessments<br>▪ Transparency to schools and/or public on school performance<br>▪ School inspections and inspections institutions |
| **Financial and organizational foundation** | ▪ **Organization structure:** The system takes steps to make the school network shape and governance manageable, and to delineate decision rights accordingly<br>▪ **Financial structure:** The system establishes an efficient and equitable funding allocation mechanism for schools | ▪ Optimization of number of schools or teachers<br>▪ Decentralizing financial and administrative rights<br>▪ Increasing funding and changing allocation model<br>▪ Organizational restructuring |
| **Pedagogical foundation** | ▪ **Learning model:** The system selects a learning model consistent with raising student capabilities, and designs the necessary supporting materials for this new model (e.g., standards, curriculum, textbooks) | ▪ School model (number of years students spend at each education level)<br>▪ Streams/tracks based on student outcomes and academic focus<br>▪ Language of instruction |

**Systems included**

| | | |
|---|---|---|
| Armenia (2003+) | Hong Kong (1983-1988) | Lithuania (1995-2000) |
| Aspire (2002-2003) | Jordan (1999+) | Poland (2000-2002) |
| Boston (2003-2005) | LBUSD (2002-2005) | Singapore (1983-1987) |
| Chile (2006+) | Latvia (1995-2000) | Slovenia (1995-2005) |

Source: McKinsey & Company interventions database and system interviews

# Exhibit 13:
# Eastern European and former Soviet states relied on the same interventions to increase school system manageability and transparency

| | Reallocate system manage-ment | Revise the school model | Optimize schools/ staff | Decentralise funding/ per-pupil funding model | Data foundations (national assessments) | Highlighted interventions |
|---|---|---|---|---|---|---|
| Armenia | ✓ | ✓ | ✓ | ✓ | ✓ | **Armenia's optimization of teachers, from 65,000 to 40,000:**<br>• Second phase of reforms (1999-present) focused on 'intra-school' optimization<br>• Minimum teacher load of 22 hrs/week mandated<br>• Rise from 9:1 student ratio in 2000 to 14:1 in 2009 |
| Latvia | ✓ | ✓ | ✓ | | | **Latvia reallocated system management roles:**<br>• State Inspectorate established to conduct school inspections (1991)<br>• State Education Centre set up for student evaluation (2004) |
| Lithuania | ✓ | ✓ | ✓ | ✓ | ✓ | **Lithuania's optimization program focused on closing small schools in order to concentrate resources within a reduced network**<br>• 1998: 2600 school<br>• 2009: 1311 schools<br>• 2012: 1000 schools planned |
| Poland | ✓ | ✓ | ✓ | ✓ | ✓ | **Poland's switch to a 6+3+3 model (from 8+4) required introducing lower secondary schools**<br>• 4000 lower secondary schools opened in one year<br>• Required shutting down and reconstituting 3764 primary schools |
| Slovenia | ✓ | ✓ | | ✓ | ✓ | **Slovenia started expanding lump sum financing to schools in 2004.** This gave schools more autonomy in distributing funds and bound them to carry out an ongoing process of self-evaluation |

Source: McKinsey & Company interventions database and system interviews

in journeys from fair to good all used strikingly similar core interventions to those adopted in Poland (Exhibit 13). This similarity is not surprising given their context: all these systems faced similar challenges in how to create and manage their national education systems following the dissolution of the Soviet power bloc; and, all at that time also had very similar student outcomes.

An important emphasis in the fair to good improvement journey stage is the introduction of system-wide student assessment systems: data plays a powerful role in this stage in two ways. First, it enables system leaders to identify whether student outcomes are improving or not and thereby allocate attention and resources to the areas of highest need. Second, it holds educators across the system accountable for raising student outcomes, helping to shift the system culture "from teaching to learning." The city of Boston and the Commonwealth of Massachusetts illustrate how these two forces combine. In 1998, Massachusetts launched the Massachusetts Comprehensive Assessment System (MCAS), a statewide tenth-grade student assessment; this became a binding graduation requirement in 2001. MCAS is judged to have among the most stringent proficiency standards of any state assessment in the United States.[15] During the 1998 MCAS pilot, roughly half of all students across the state failed the assessment. In 2001, at the point MCAS became binding on the state, Massachusetts used the test results to allocate resources to the neediest districts. Of the approximately USD 55 million in statewide funding that followed the first binding MCAS in 2001, USD 5 million went to Boston to fund double-block classes (whereby students stay in the same class for two periods in a row), summer programs, and after-school programs. Massachusetts also used the 1998 pilot data as the funding rationale for a professional development program for 1,000 urban principals in 2001. Starting from an initial 40 percent pass rate at their first sitting of MCAS in 2001, the class of 2003 achieved an 80 percent pass rate by the time they were twelfth-graders. According to one Boston leader from the early years of the program, "Without the additional resources for the class of 2003, we would not have gotten the improved results."

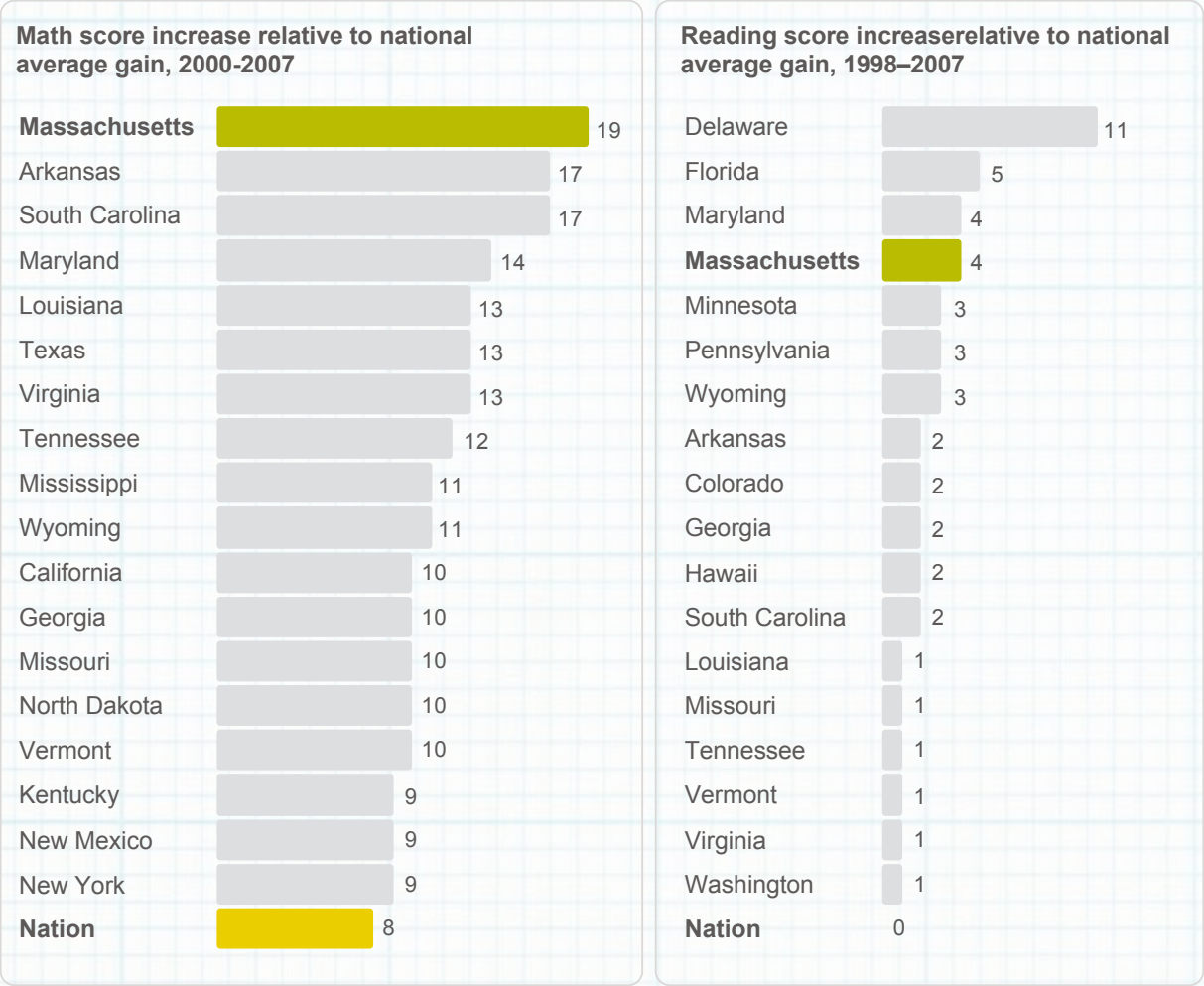
To support its schools in achieving higher outcomes, the city of Boston created the MyBPS data system. This contained detailed student achievement data accessible to teachers, principals, and administrators. Boston's district leaders reviewed this data and invited teachers with track records of demonstrated success to speak to the leadership about their teaching or to contribute to teacher study groups. Yearly targets were set for each school for increasing their student achievement levels and for closing any achievement gaps between socioeconomic sub-groups. Schools that were performing well were allowed more flexibility; those that performed poorly received greater intervention from the district office. This pattern of interventions is seen across systems on the fair to good journey; for example, England called this, "intervention in inverse proportion to success". Massachusetts was able to take intervention further than most. The state had succeeded in removing its principals from collective bargaining, so the district held its principals accountable for their school's performance. During Tom Payzant's eleven-year tenure as Superintendent of Boston Public Schools, 75 percent of all the district's principals were either replaced or retired.

Between 1998 and 2007, Massachusetts registered the highest gains in the United States on the National Assessment of Educational Progress (NAEP), making the largest gains in math and the third-largest gains in reading of all U.S. states (Exhibit 14). By 2007, it was the top-performing state in the U.S. on both NAEP's reading and math assessments. Within this much-improved state, the Boston Public School District is a much-improved district. As a four-time finalist and 2006 winner of the Broad Prize for Urban Education, Boston has raised the proportion of its students that pass the state exams in mathematics from 23 percent in 1998 to 84 percent in 2008, and those that pass in reading from 43 percent in 1998 to 91 percent in 2008.

The systems examined here, all of which are undergoing the journey from fair to good, show two distinctive but overlapping sets of objectives. The first group comprises the countries from Eastern Europe that only recently emerged from under communism; these systems focused on ➔

# Exhibit 14: Massachusetts was the most improved US state on NAEP during 1998-2007

**Systems with above average increases in NAEP scores, 8th grade**




2007 National average[1] 8th grade mathematics score was 280

2007 National average[1] 8th grade reading score was 263

Source: National Assessment of Education Progress (NAEP)

# Exhibit 15: Good to great journeys emphasize shaping the professional

| Theme | Description | Example interventions |
|---|---|---|
| **Raising calibre of entering teachers and principals** | ▪ **Recruiting:** The system raises the entry bar for new teacher candidates<br>▪ **Preparation and induction:** The system raises pre-service training quality and certification requirements | ▪ Recruiting programs<br>▪ Pre-service training<br>▪ Certification requirements |
| **Raising calibre of existing teachers and principals** | ▪ **Professional development:** The system raises professional development requirements and provides more opportunities for self-, peer-, and center-led learning and development<br>▪ **Coaching on practice:** Instructional coaches work with teachers to strengthen their skills in areas such as lesson planning, student data analysis, and in-class pedagogy<br>▪ **Career pathways:** The system creates teacher and leadership specializations through career pathways, raising expectations with each successive pathway rung and increasing pay accordingly | ▪ In-service training programs<br>▪ School-based coaching<br>▪ Career tracks<br>▪ Teacher community forums |
| **School-based decision-making** | ▪ **Self-evaluation:** The systems cultivates ownership in schools for improvement through introducing self-evaluation for schools and making performance data more available<br>▪ **Flexibility:** The system gives schools the flexibility to pursue specialized programs appropriate to their students, and increasingly decentralizes pedagogical rights | ▪ Self-evaluation<br>▪ Data systems<br>▪ Independent and specialized schools |

| **Systems included** | | | |
|---|---|---|---|
| | Aspire (2003+) | Long Beach (2005+) | Saxony (2000-2005) |
| | Boston (2006+) | Latvia (2001+) | Singapore (1988-1998) |
| | England (1995+) | Lithuania (2001+) | Slovenia (2006+) |
| | Hong Kong (1989-1999) | Poland (2003+) | South Korea (1983-1998) |

Source: McKinsey & Company interventions database and system interviews

reshaping and optimizing their system management. The second group, that has not had to face the burden of undergoing nationwide structural change, nevertheless focused on introducing system-wide performance management and assessment systems. This again underlines the pattern we see in how systems at the same performance stage, whatever their context, draw from the same performance objectives and use the same intervention cluster in addressing these objectives.

## The "good to great" journey: shaping the teaching profession

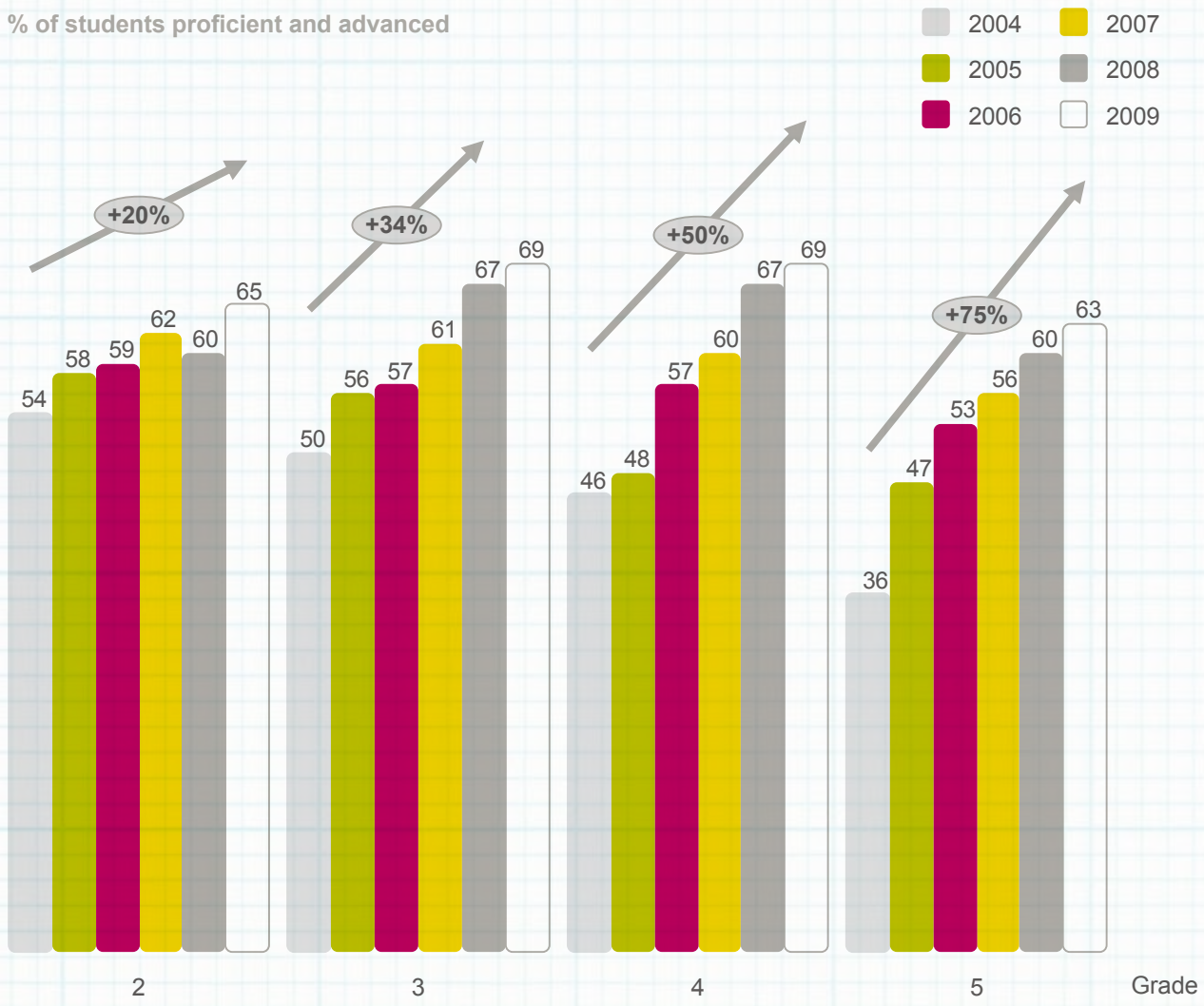
Once the foundations are in place, in the next stage of its journey the system turns its attention to the professionalization of its educators. The path to school system improvement now relies on the fidelity of educators' practice in their teaching and learning routines. Whereas the success of previous improvement journey stages largely relied on central control over the system and its educators, the good to great journey marks the point at which the school system comes to largely rely upon the values and behaviors of its educators to propel continuing improvement. To this end, in systems on the good to great journey, the center employs a cluster of interventions aimed to make the apprenticeship and mentorship of educators as distinct as that seen in other professionals such as medicine or law (Exhibit 15).

Long Beach Union School District (LBUSD) in California provides an example of the development of these routines and practices. An LBUSD leader described their aspiration for professionalization as follows: "We wanted all our educators to speak a common language about the craft of teaching, and to have the same calibration of what quality teaching and learning looks like . . . Our litmus is would you put your child in this school?" Indeed, in interviewing over fifteen system leaders across LBUSD, the mantra of "would you put your child in this school" was echoed in nearly every discussion.

LBUSD engaged in multiple interventions to achieve this goal. "Our starting point is always looking at the kids and looking at the data," says one system leader. Driven by the ethos that data creates objectivity in decision-making, student performance data (test grades, homework assignments) is available throughout the system on "School Loop"; all stakeholders, including parents, have access to it. This data transparency is paired with "walk-throughs," whereby the superintendents at each level (primary, intermediate, secondary) walk through the schools and classrooms with principals, coaches, and others to discuss the data and the school goals. In the case of struggling schools, there can be several walk-throughs with the principal each month. One system leader says, "Walk-throughs must be respectful and unifying, but they also open up the school to review. We look at the data knee-to-knee with the principal, we listen, we ask questions, we give feedback on how the data relates to the school goals, and we give praise where warranted." A walk-through may sometimes involve principals from other schools with similar learning objectives. It is worthy of note that the spirit of LBUSD's walk-throughs is analogous to the weekly "grand rounds" in medical teaching, where medical peers present the patient case, ask questions, explore alternatives, make a diagnosis, and develop a treatment plan.

On the basis of LBUSD's walk-throughs and the School Loop data, the district allocates its coaching resources to support struggling schools. It has created specialized curriculum coaches for its teachers: these are expert teachers in priority areas (math, literacy, and college-readiness) who are assigned to four or five schools, and who generally work with three teachers in a school on any given day. They coach teachers in a three-step sequence of "see one, share one, do one," whereby the coaches first run a demonstration class, then co-teach a class with the class teacher, and finally observe the teacher instruct the class alone. This "gradual release" spans a period of three to four weeks. This sequence is again coupled with walk-throughs, whereby the coach and principal walk through the classes of the teachers being coached, with the coach providing the principal with guiding questions. "Coaching must be linked to principals so that they can follow-up," says one LBUSD coach. Differing significantly from the scripted coaching in the poor to fair journey, coaching in the good to great journey focuses on the transmission of effective teaching strategies rather than "tips and tricks." Great care is taken to ensure that the teacher remains empowered throughout the coaching process – for example, a coach only ➔

# Exhibit 16:
# Long Beach math scores on the California STAR examinations improved significantly between 2004-2009




Source: Long Beach Unified School District

enters the classroom with the teacher's permission and the coach maintains full confidentiality with the teachers with whom they are working. For similar reasons, the coach will not comment about the teacher's performance to the principal, for instance, but instead suggest a walk-through so that the principal can see the situation for herself. At all times, the coaches ensure that it is clear that the teacher and principal remain responsible for instructional quality; the coach only supports them in delivering this. Similarly, for principals, LBUSD has created intervention program coaches. These are former high-performing principals who spend two to three hours each week with each of the principals of struggling schools who are undergoing coaching.

LBUSD has also undertaken multiple interventions to ensure that its new teachers are inculcated in what they call "the Long Beach way." LBUSD recruits eighty percent of its teachers from the School of Education at California State University at Long Beach, and so has sought to train prospective teachers in the instructional practices used by the district from the outset. To this end, staff from the LBUSD curriculum department teach the classes on teaching method at the School of Education. "In 1994, there was little interaction between the university and LBUSD. Now it is at the point where you can't tell the difference between who is from the district and who is from the university," says a Cal State leader.

LBUSD provides all its newly trained teachers with three to five hours of coaching a week during their first year in their first school in order to embed good instructional practice and classroom management skills. The coaching is provided by a trained senior teacher from the teacher's own school. During years two to three, teachers receive seven days of professional development training each year. "Teachers were being pulled apart by lots of training from different silos ... We centralized the training and made it more coherent," observed one LBUSD leader.

Lastly, when teachers register impressive student gains, LBUSD is proactive in noting and understanding their practices. Along the same lines as "evidence-based medicine," it identifies the best delivery methods from pilot data and then rolls out the program across all its primary schools. For example, a math teacher developed a new program in his classroom for primary math instruction, known as the MAP^2D program. Inspired by his aunt, who taught math in Singapore, this high-performing math teacher's program prescribes a specific lesson structure that improved his students' proficiency from 40 percent to 60-70 percent. LBUSD allocated four math coaches to work with the math teacher to codify and pilot the program in other schools.[16] LBUSD's results in the California STAR examinations show it has achieved 20-75 percent improvement in grades two to five during 2004-09 (Exhibit 16)

LBUSD's example shows how systems further along the performance journey start to focus on reinforcing the pedagogical aspects of teacher learning and performance. The energy for this process also starts to shift away from the center to the schools themselves; we will talk more about these aspects of school system improvement in Chapter 3, "Sustaining."

### The "great to excellent" journey: improving through peer-led support and teaching innovation

In the final frontier of school improvement, the journey from great to excellent, systems focus on creating an environment that will unleash the creativity and innovation of its educators and other stakeholder groups. At this point in the improvement journey, system educators are highly skilled and have a body of agreed routines and practices that have become innate to how they work. The intervention cluster for the journey from great to excellent serves further to enhance the educators' responsibility for looking after each other's development; the systems give their teachers the time, resources, and flexibility to reflect upon and try out new ideas to better support student learning (Exhibit 17).[17]

The systems that have embarked upon the journey from great to excellent have each chosen different approached to sparking innovation amongst their educators. For example, Hong Kong created "The Quality Education Fund," an endowment of HKD 50 billion to support schools that undertake approved school improvement projects or ➔

# Exhibit 17: Great to excellent journeys emphasize learning through peers and innovation

| Theme | Description | Example interventions |
|---|---|---|
| **Raising calibre of entering teachers and principals** | ▪ **Learning communities:** The system facilities school-based learning communities to create peer-led support and accountability to each other<br>▪ **Flexibility:** The system provides effective educators with greater pedagogical autonomy<br>▪ **Rotations:** The system rotates educators throughout the system in order to spread learning and varied styles of mentorship | ▪ Collaborative practice amongst educators<br>▪ Decentralizing pedagogical rights to schools and teachers<br>▪ Creating rotation and secondment programs across schools, and between the center and schools |
| **Creating additional support mechanisms for professionals** | ▪ **Leverage:** The system provides administrative staff in schools so that teachers and principals can focus on pedagogy and leadership rather than administrative tasks | ▪ Providing additional administrative staff |
| **System-sponsored innovation across schools** | ▪ **Stakeholder innovation:** The system sponsors and identifies examples of innovative practices in schools (teaching and learning practice, parent/community involvement practices, etc.) and then develops mechanisms to shares these innovations across all schools | ▪ Sharing innovation from the front-line<br>▪ Funding for innovation |

**Systems included[1]**

Hong Kong (2000+)
South Korea (1999+)
Ontario (2003+)
Saxony (2006+)
Singapore (1999+)

1 All these systems are on the journey from Great to Excellent and, while they demonstrate rising gains, none have yet attained the Excellent threshold on the universal scale (560+)

Source: McKinsey & Company interventions database and system interviews

action research. Such research is meant to be highly practical, with immediate benefit to teaching and learning practices. Similarly, South Korea also funds action research by teachers and counts these efforts toward their professional development requirements. Districts make grants available to schools that lead their own research projects – each school can select a research topic, conduct research, publish the results and invite teachers from other schools to peer-review their findings. One South Korean principal estimates that 150 schools out of 1,000 schools in his province have conducted such research. The districts also fund inter-school learning, whereby teachers from a number of different schools in a district can apply to jointly conduct research spanning all their schools. Participation in all types of research is an important consideration in South Korea's annual teacher reviews, incentive allocation, and promotions. In pursuit of this same theme of "making practice public," schools encourage teachers to open up their classrooms to others two or three times a month, at which times other teachers can come and visit and observe their lessons.

## Prescribe adequacy, unleash greatness

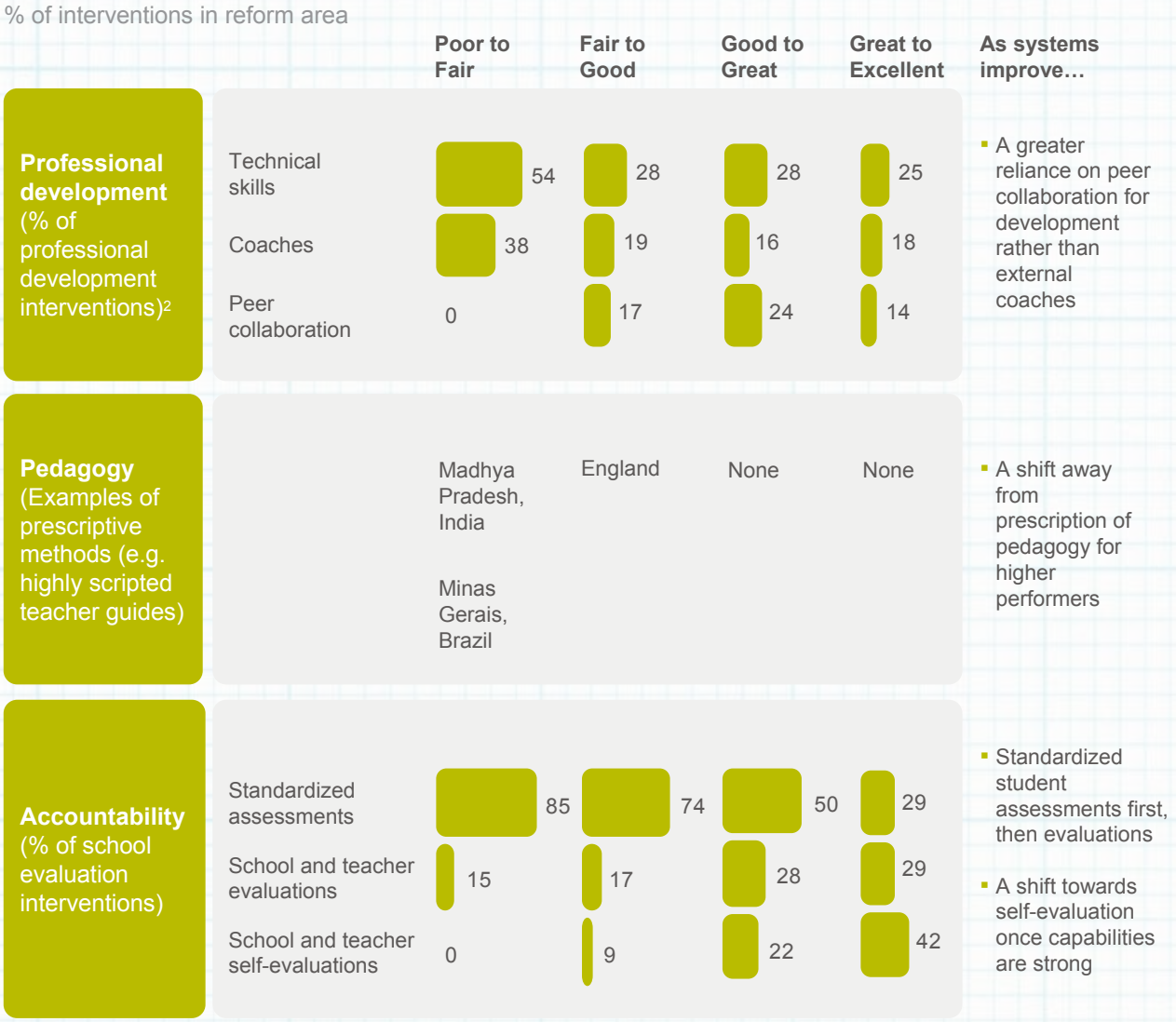
Implicit in the previous discussion of the intervention clusters associated with each improvement journey stage, is the observation that a striking correlation between a system's performance level and the tightness of the central control exerted on schools across all the improvement journey stages in our sample. Systems with teachers that are less grounded in the system pedagogy exercise tighter control over teaching and learning processes, while systems with higher educator skills loosen it. This is not to say that teachers "make it up on the fly" in their classrooms, but rather that the system is sufficiently developed that it has already ensured evidence-informed and school-based instructional practice whereby teachers collaborate and set standards to which they hold each other accountable.

Exhibit 18 shows the dominant pattern emerging from our sample systems. As system performance rises, professional development shifts away from a focus on technical training delivered by central coaches to a greater reliance on teacher-peer collaboration and development. As the level where system performance is above the global average, there is no longer a single example of a system prescribing scripted instruction in our sample. As system performance rises, accountability also expands, moving from center-led standardized student assessments to also include school and teacher self-evaluation. In order to achieve improvement in student outcomes, lower-performing systems focus on raising the floor, while higher performing ones focus on opening up the ceiling.

There is therefore a correlation between the improvement journey stage and the form the intervention takes. Why is this the case? One answer is that the main challenge of systems engaged in the poor to fair and fair to good stages is to minimize performance variation between classes and across schools. This requires ensuring that lower-skill teachers are given the support of high-quality teaching materials and lesson plans that can closely guide what they do on a daily basis. As one Asian system leader says looking back at his system's poor to fair journey, "We did everything we could to make it as easy as possible for our teachers to teach." However, when teachers achieve a higher level of skill, as is the case in good to great and great to excellent improvement journey stages, such tight central control becomes counterproductive to system improvement. Rather, school-level flexibility and teacher collaboration become the drivers of improvement because they lead to innovations in teaching and learning. The center learns from these school-based innovations and then encourages their use in other schools across the system. Higher skill teachers require flexibility and latitude in how they teach in order to engage in such innovation and to feel motivated and fulfilled as professionals. Exhibit 19 illustrates the pattern of how our sample of improving systems increasingly decentralized pedagogical rights, defined as choosing methods and materials for instruction and curriculum, to the middle layer or schools as performance improved.[18]

The contrast between the approaches used by Madhya Pradesh (India) and Ontario (Canada) in boosting student outcomes is illustrative of ➔

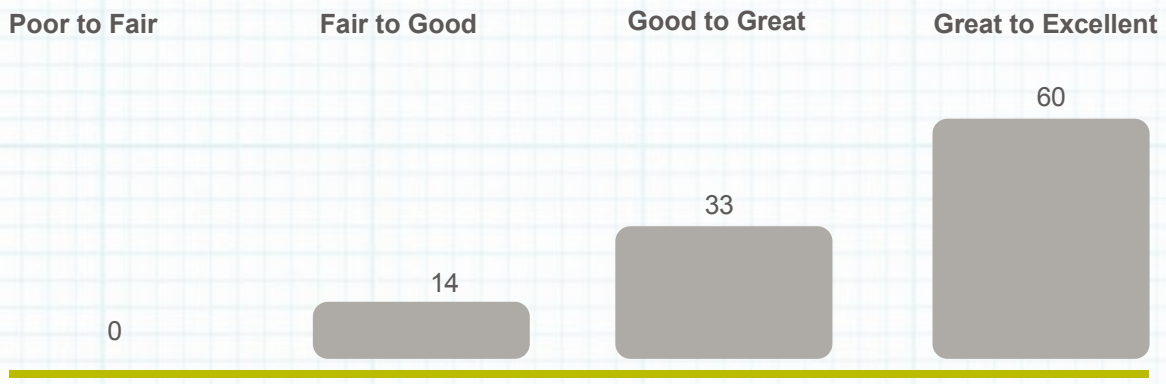
# Exhibit 18:
# There is a shift from central guidance to school-based collaboration and self-evaluation as performance levels increase




1 Sample sizes (Poor to Fair, Fair to Good to Great, Great to Excellent). Professional Development and Training (N=13, N=28, N=36, N=25). Accountability (N=13, N=23, N=18, N=7)
2 Other types of professional development and training interventions not listed here account for 33% of all professional development and traing interventions. They include career tracks; pre-service recruitment certification, and training; and job rotations
Source: McKinsey & Company interventions database and system interviews

## Exhibit 19:
## Across our sample systems, the ‘center’ increasingly decemtralised pedagogical rights as performance increased

**Pedagogical rights**
% of systems in reform phase that decentralized pedagogical rights to middle layer or schools




Source: McKinsey & Company interventions database

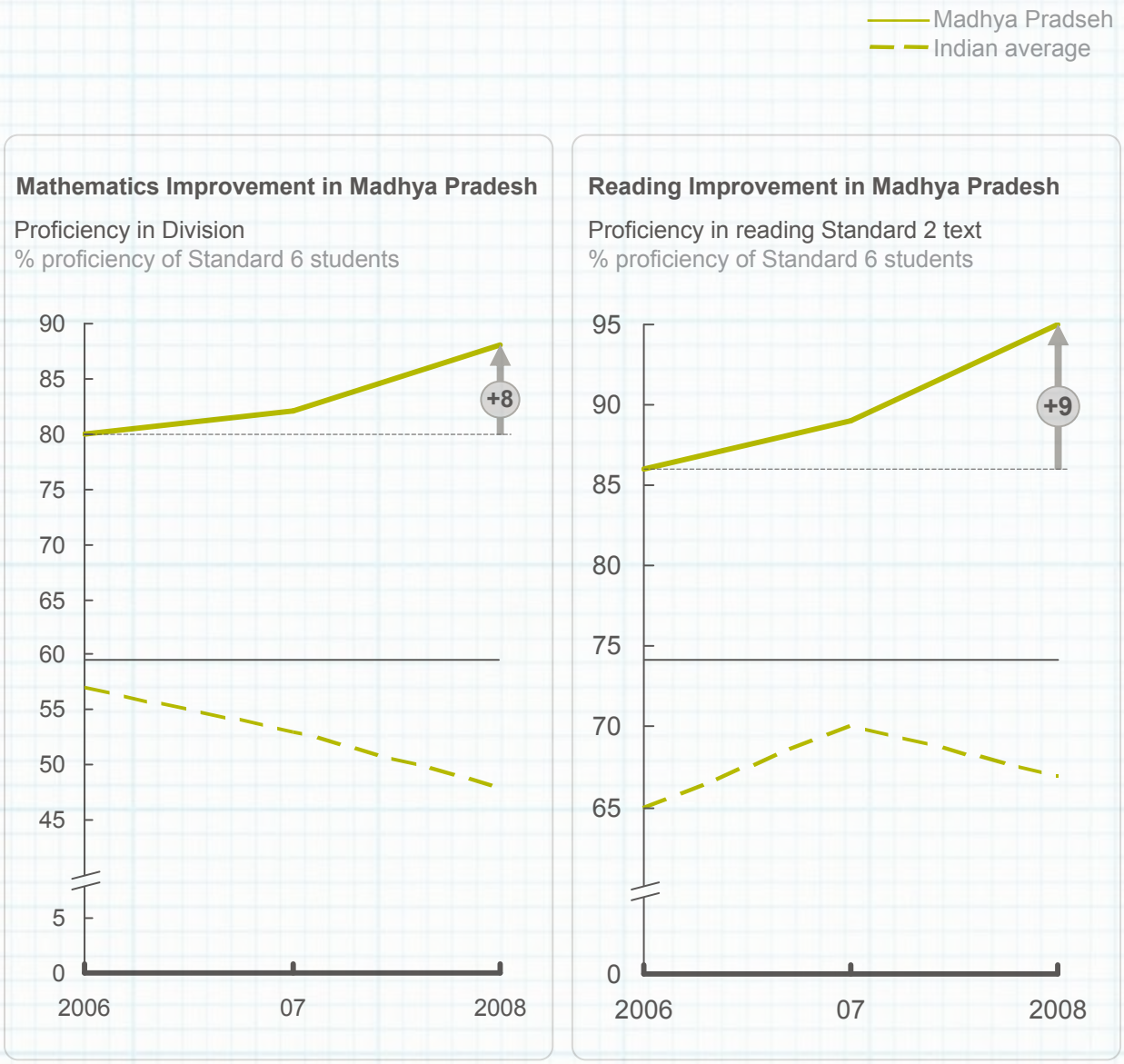
this pattern. With over 138,000 schools, Madhya Pradesh is the largest school system we studied. It also had among the lowest student achievement levels of all the improving systems at the start of its reform journey. Less than 50 percent of standard one and two students (ages six to seven) could read individual letters and words, as compared to the national average of 70 percent; less than 60 percent of its students in standards three to five (ages eight to ten) could read a standard one text, as compared to the national average of 67 percent. At this point in its journey, the system lacked standards for students learning outcomes, rigor in test assessments, and accountability in its schools.

In 2005, Madhya Pradesh launched a literacy reform called "Learn to Read," using a highly standardized teaching model designed to improve teaching and learning throughout the school system. Every classroom lesson now had a newly prescribed lesson plan, teaching materials, student worksheets, and a set of teaching techniques based on activity-based learning. A central group of 75 skilled trainers cascaded these teaching practices throughout each district, block, cluster, and school throughout Madhya Pradesh using a "train the trainer" model. Student comprehension was assessed once a month using standardized tests, and the center closely tracked learning data to assess the progress of each district, school, and child, such that any deviation was immediately detected. The cascade of training was supplemented by larger annual ten-day training programs for teachers; there were monthly refresher training sessions for all teachers delivered by satellite broadcast; in-class observation of teaching was carried out by cluster managers, bloc leaders, and district leaders; 350 teacher coaches provided targeted support to areas of need identified from students learning data. The collective intent of all these interventions was to drive uniformity in classroom practice throughout the system. As one system leader described, "Our standardization is so comprehensive that if a student in a class in one corner of the state is put into another school in a totally different location, he would not even notice the difference ... Everyone has to teach the same curriculum at the same time in the same way." In the period 2006-08, this large-scale standardized model, with its tight controls and monitoring resulted in significant quality improvement. The proportion of Madhya Pradesh's standard six students who could read a standard two text increased from 86 to 95 percent, far exceeding India's national average performance during the same period (Exhibit 20).

We observe similar examples of increased central guidance and oversight across the improving systems undergoing poor to fair journeys. Scripted teaching practices, proficiency targets for each school, frequent, standardized testing to monitor system progress, training cascaded from the centre for all teachers, and the provision of external coaches all typify the interventions made on these journeys, from Minas Gerais to Mahdya Pradesh.

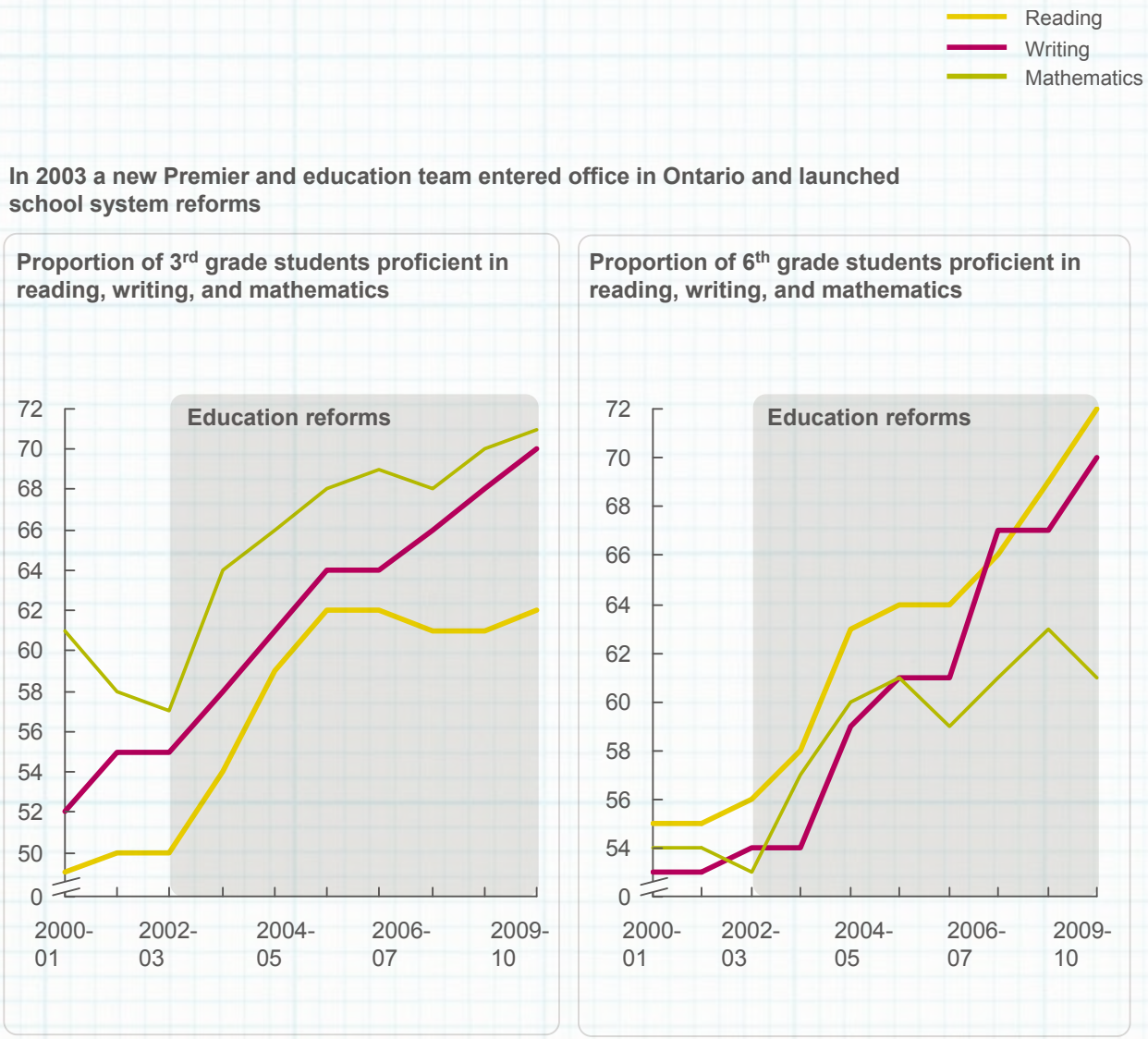
The conditions faced by Ontario are very different to those in Madhya Pradesh. Ontario, which also has a relatively large school system of nearly 5,000 schools, 120,000 teachers, and 2.2 million students, is among the world's highest-performing school systems. It consistently achieves top-quartile mathematics scores and top-decile reading scores in PISA. The system is not without its challenges, however, so Ontario's leadership was seeking to further improve its system performance. Its own provincial assessments rated over half of its third and sixth-graders as below standard in reading, writing, and mathematics; the school system had lost 26 million student-days over the previous four years due to work stoppages; as a result, public confidence in the school system was low and families were increasingly choosing to exit the public system. In consequence, in 2003, Ontario's new leadership developed a strategy to improve primary literacy and numeracy, and to raise secondary completion levels. Like Madhya Pradesh, Ontario focused on primary literacy. Recognizing that improved teaching practices was the key improvement driver, it set an ambitious improvement aspiration, and tracked student-learning data regularly to assess progress and direct support accordingly. However, this is where the similarities between the two systems end. Unlike Madhya Pradesh, Ontario did not centrally script and cascade new teaching and learning practices to all classrooms. Instead, it focused on cultivating school-led innovation and improvement. As one Ontario system leader described, "We minimized the amount of directing or ➔

## Exhibit 20:
## Madhya Pradesh both outperformed and improved more than the overall Indian average in mathematics and reading over two years




Source: Madhya Pradesh education department

# Exhibit 21:
# Ontario's proficiency levels show consistent improvement at both 3rd and 6th grades




Source: Ontario Education Quality and Accountability Office; IELD Ontario Case Study Report 2007

mandating we did. Instead, we needed methods to get school professionals' ideas so we could build on them. We regularly brought people together to share their practices and exchange ideas. We did almost no mandating of specific strategies – we got them to develop their own plans. We didn't micromanage schools or districts in this process. We empowered them." Between the period of 2003 and 2007, Ontario registered a strong rise in student proficiency levels in both third and sixth grades (Exhibit 21). Following a slight plateau in 2007, Ontario's rises have continued through 2009/10 for four of the six assessments. This pattern of a steep rise followed by a plateau is a common phenomenon (England and Boston have also experienced it). This might be due the fact that once the "easy wins" have been achieved in classroom instruction, further improvements take longer to embed and are harder to achieve.

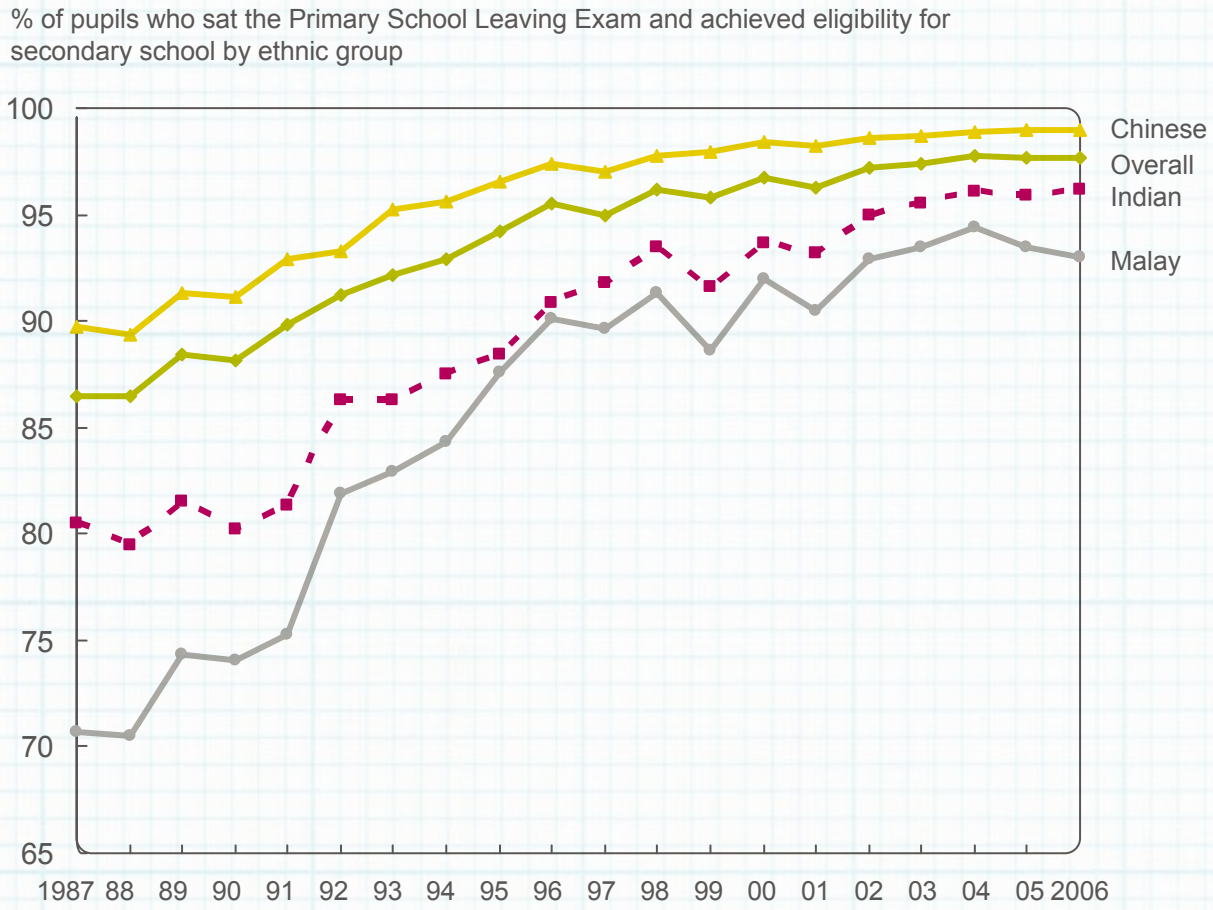
Singapore provides an example of how a system shifts in emphasis as it goes through the various stages of the entire improvement journey, from poor to great, as Singapore as done over the past forty years. During this time it has decreased central guidance on teaching and learning as its system performance has risen. Singapore system leaders describe their system as having gone through three phases: "Survival-driven" (1959-78), "Efficiency-driven" (1979-96), and "Ability-driven" (1997-present).

Singapore's Survival phase was primarily focused on enrolment and ensuring that every child had a school seat. This resulted in schools being built at the rate of one per month and the teaching force doubling, from 10,500 in 1959 to over 19,000 by 1968. By the end of this period, Singapore had achieved near universal primary education. However, almost thirty percent of primary school pupils did not progress to secondary school, and English language proficiency was low and educational wastage high (in terms of failing to achieve the expected standards and leaving school prematurely).

The Efficiency phase focused on reducing performance variation across the school system. "Our challenge was how to achieve above average outcomes from below average inputs," recalls one Singaporean leader. Students were streamed into different tracks based on their aptitude, not only to reduce dropout rates, but just as importantly, to ease the burden on teachers so that they only taught classes of students with similar capability levels. Simultaneously, Singapore created the Curriculum Development Institute of Singapore (CDIS) in 1980 in order to develop a suite of supporting teaching materials that could be used off-the-shelf by less-experienced and less-skilled teachers. A Singaporean system leader recalls, "For each lesson, we created the lesson plan, the teacher manual, the student workbook, and the activity or experiment or video that would open the lesson." Each classroom in the same grade and subject level received exactly the same resources, and CDIS held workshops with teachers to explain how to use the materials effectively. Moreover, teachers had to keep a record book of their classroom activities, which were submitted to the principal every Monday. Regular student assessments enabled the Ministry of Education to monitor student outcome progress. As one system leader noted, "We were highly prescriptive in our teaching and had a mass production mindset ... We were textbook-bound and examination-driven." Through the 1980s and 1990s, Singapore raised the floor of performance in the system significantly, and narrowed the achievement gap across its ethnic groups (Exhibit 22).

Singapore moved from rigid prescription to greater flexibility as it embarked on its good to great improvement journey. By the end of the 1980s, Singapore had introduced school formats that had greater autonomy, including establishing Independent Schools in 1988 and Autonomous Schools in 1994. By 1995, Singapore's school system was among the top-performing systems in the world, topping TIMSS rankings in both math and science that year. The Curriculum Development Institute of Singapore closed its doors in 1996 because "it was no longer needed." Then, in 1997, Singapore launched "Thinking Schools, Learning Nation" (TSLN), marking the start of its Ability phase and emphasizing a shift in focus toward enabling each student to reach the maximum of his or her potential. This focus on student ability required schools to be given much greater flexibility and responsibility for how they should teach and manage their students. ➔

# Exhibit 22:
# Singapore narrowed the achievement gap between the ethnic groups




Source: Singapore Ministry of Education

TSLN gave teachers greater freedom in classroom practice, and gave principals decision rights on school management matters. It introduced school clusters[19] to create a peer-based forum for school leadership development and the sharing of effective teaching and learning practices across schools. It also changed its school inspection model, replacing the previous highly centralized model with a more collaborative one focused on self-assessment and quality assurance.

Throughout the latter period, Singapore worked intensively on strengthening the caliber of its teachers and principals so that they could make the best use of their greater freedoms. It established a system that accommodated three career tracks (Leadership, Teaching, and Senior Specialist), narrowed recruitment into teaching to the top one-third of each graduating cohort, expanded professional development to one hundred hours per year, and creating mentorship pairings for school leaders. More recently it has focused on strengthening the networks of Professional Learning Communities (PLCs) in schools that encourage teachers to collaborate with one other in reviewing and improving their classroom practice. In the words of one system leader, "As the skills of our educators rose, we needed to change our approach in how we managed them. We could no longer prescribe what they did, we had to treat them like professionals who had good judgment, knew their students well, and who could make their own decisions."

Singapore's experience focuses us on a second important question about control: whether the degree of prescription and flexibility within a school system can or should vary within that system. The answer is that it can: there are examples in our sample in which the school system has given more attention to scripting its low-performing schools while providing more flexibility to the higher performing ones. England, when on the fair to good journey in 1996, launched a literacy and numeracy program, tightening guidance over classroom practice. The system cascaded improvement targets to each school and mandated that every day every student should be given a "literacy hour" and a "daily maths lesson," and provided teachers with models of literacy and numeracy instruction that they were expected to use if they could not demonstrate higher performance on their own. High-performing schools that were already exceeding minimum proficiency targets were exempt from this prescription. Similarly, Western Cape, during its poor to fair journey, gave higher performing schools more freedom; while lower-performing schools were required to follow the guided literacy programme, higher-performing schools were exempt. In the words of one WCED system leader, "We cannot support all our schools deeply. If a school is doing well — in the seventies or more — we really pay very little attention to it except to learn what it is doing. They have more freedoms."

This correlation between system performance and the degree of central control over the school system has parallels to lean operations. A given production system must combine inputs and process in order to produce output. When input quality is low, the production system must have tight processes in order to deliver a quality output. Conversely, when input quality is high, the production system can loosen the processes to produce the same output. More specifically, across all sectors to which lean operations approaches are applied, there are two steps: first one must stabilise the system, then second, shift it into continuous improvement. The objective of stabilising is to quickly arrest a system in crisis and achieve an immediate step-change in performance to a uniform adequate level. Intervention at this stage requires tighter central process control, with scripted standard operating procedures, "back to basics" simplification of production processes, the creation of reliable data on system performance, tighter governance, such as regular reporting and performance reviews, and re-establishing a shared sense of purpose that is cascaded through all levels of the system. All of these interventions reflect the observations we made of school systems that have improved from poor to fair to good performance levels.

Once stabilised, systems then shift to continuous improvement. In contrast to stabilising, the objective of continuous improvement is to build ever-higher excellence through cultivating consistent and incremental frontline-led improvement. Frontline managers are empowered as agents of change, with

daily team huddles, feedback sessions, and formal mechanisms for the system to collect, evaluate, and disseminate innovation that occurs in the front line. As a result, when a breakthrough is achieved, it quickly sets a new standard to be maintained across the system. Talent development also becomes more collaborative and based on apprenticeship. For example, many lean systems introduce "quality circles" as forums for talent development. More robust performance management systems also become important for creating transparency throughout the system, and enable quick identification of upward or downward spikes in performance. These performance management systems are reinforced by incentives that are relevant to the sector and job functions within it. These interventions mirror the observations we made of school systems that have moved from good towards excellence.

The high degree of resonance between our observations of school system improvement and lean operations raises the potential for school systems to learn from lean/ systems. Our intent here is not to imply that school systems are manufacturing processes but to indicate and appreciate that systems in other domains also alter their processes based on their input characteristics. The data pattern of our research shows that our sample school systems fully recognized the caliber and needs of their teachers and principals at each performance journey stage, and varied their approach accordingly, in order to achieve improvement.

## Common but different

We have identified a group of six interventions that occur with equal frequency across all performance journeys, but manifest themselves differently at each improvement journey stage. The stability of these six interventions is unsurprising given their centrality to teaching and learning:

- Technical skill building: strengthening professional development for new and tenured teachers and principals.
- Student assessment: assessing students at the regional or national level for various grades and subjects.
- Data systems: gathering, analyzing, and sharing data on system performance (schools, students, educators, geographic areas), and using data as a tool to direct the allocation of system support.
- Revised standards and curriculum: defining what students should know, understand, and be able to do, and creating the accompanying teaching content.
- Teacher and principal compensation: introducing a reward schemes for high performance, and structuring teacher and principal compensation in accordance with the role they play.
- Policy documents and education laws: facilitating the improvement journey by articulating the aspirations, objectives, and priorities of the reform program.

Taking the example of teacher and principal compensation, the overwhelming majority of our sample systems ensure they have an appropriate reward and remuneration system in place for the level of skills of their principals and teachers, raising salaries where necessary and/or introducing rewards. Across all the performance journeys, system leaders were quick to note that salaries were only increased once the system had made significant progress in achieving the goals of that stage's intervention cluster. ➔

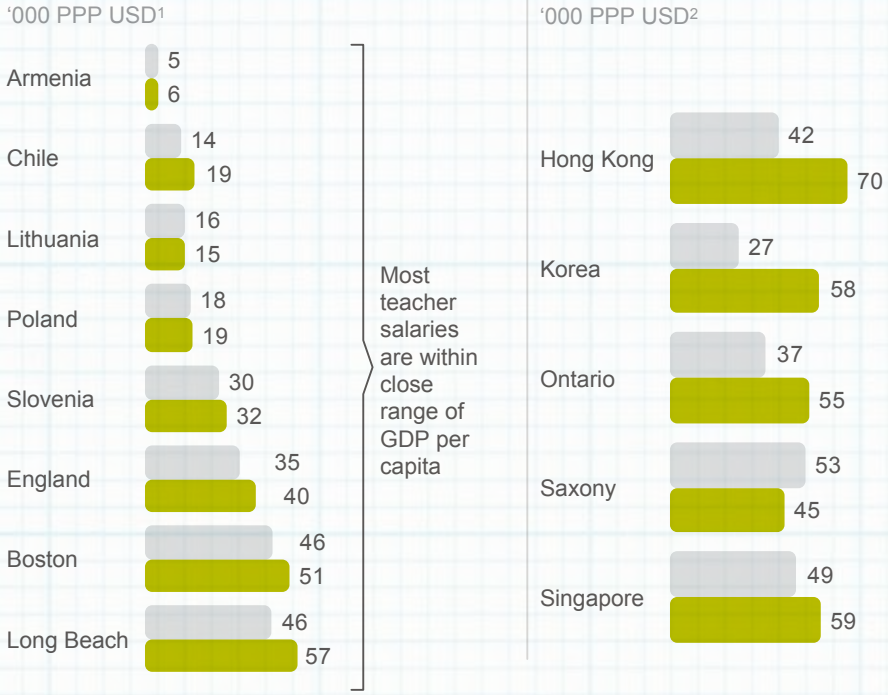
# Exhibit 23:
# Changes to teacher compensation structure varied by improvement stage

| Poor to fair | Fair to good, Good to Great | Great to Excellent |
|---|---|---|
| **Incentives funding**<br>▪ Rewards to either individual teachers or the school for achieving high performance against student proficiency targets | **Achieving 'respectable' salary**<br>▪ Base salaries at a level comparable to GDP/capita | **Attracting top talent**<br>▪ Teachers base salary significantly above GDP/capita |

**Western Cape**
- Cash prizes of ~2,000 USD given to the best performing schools in each wealth quintile

**Madhya Pradesh**
- Teachers awarded Rs. 5000 (equivalent to an additional month of salary) if they achieve and sustain student proficiency targets over 6 months

**Minas Gerais**
- Teachers received up to 1 extra monthly salary per year based on school achievements




1 Chile – average between starting and 20 year experience salary, 2007; Armenia 2009; Poland – average between trainee, contract, appointed and teacher with a diploma, September 2009; Lithuania – NET teachers' salary; Slovenia – average salary, 2008; USA GDP/Capita used for Boston and Long Beach; England 2007 data.

2 Average salary for Graduate Master/Mistress used for Hong Kong; average between starting, experienced and top salary used for Korea; average salary of teacher used for Saxony; graduate teacher with 5 years of experience salary used for Singapore; average salary of workers in Ontario schools; Canada GDP/Capita used for Ontario; Saxony-specific GDP/capita figure with Germany PPP conversion

Source: IMF; OECD; Saxony Statistical Office; Master pay scale Hong Kong; Statistical bureau of Slovenia; Lithuania department of statistics; Ministry of Education, Singapore; Interviews; UK Department of Education; Penn World Table; public dates, McKinsey & Company interventions database

# Exhibit 24: Process is the most prevelant intervention type relative to structure and resource

| Intervention type | Descriptions | Share of interventions Percent, 100%=573 |
|---|---|---|
| **Processes –** practices, activities, rights and responsibilities in the system | ▪ Change the content of **what** the system delivers<br>– Introduce standards and teaching materials<br>– Curriculum and textbooks<br>▪ Improve **how** the system, including its people, delivers content:<br>– Improve pre-service training & in-service development<br>– Raise leadership capacity<br>– Establish accountability mechanisms (assessments and targets, school inspections/evaluation, staff appraisal/promotion) | 70 |
| **Structures –** organizational, financial, and instructional configuration/shape of the system | ▪ Organizational configuration:<br>– Introduction of a 'middle' layer<br>– New institutions (e.g., assessment agency)<br>▪ Financial configuration:<br>– Optimization of the number of schools in the system<br>– School choice, or vouchers/privatization<br>▪ Instructional configuration:<br>– Student streaming<br>– School years and levels (e.g. 6+3+3 to 4+4+4) | 15 |
| **Resources –** Level and allocation of financial and human resources to fuel the system | ▪ The number and entering-caliber of people (teachers, principals, staff)<br>▪ The capacity and quality of infrastructure<br>▪ Money for reforms (e.g. raising teacher salaries) | 15 |

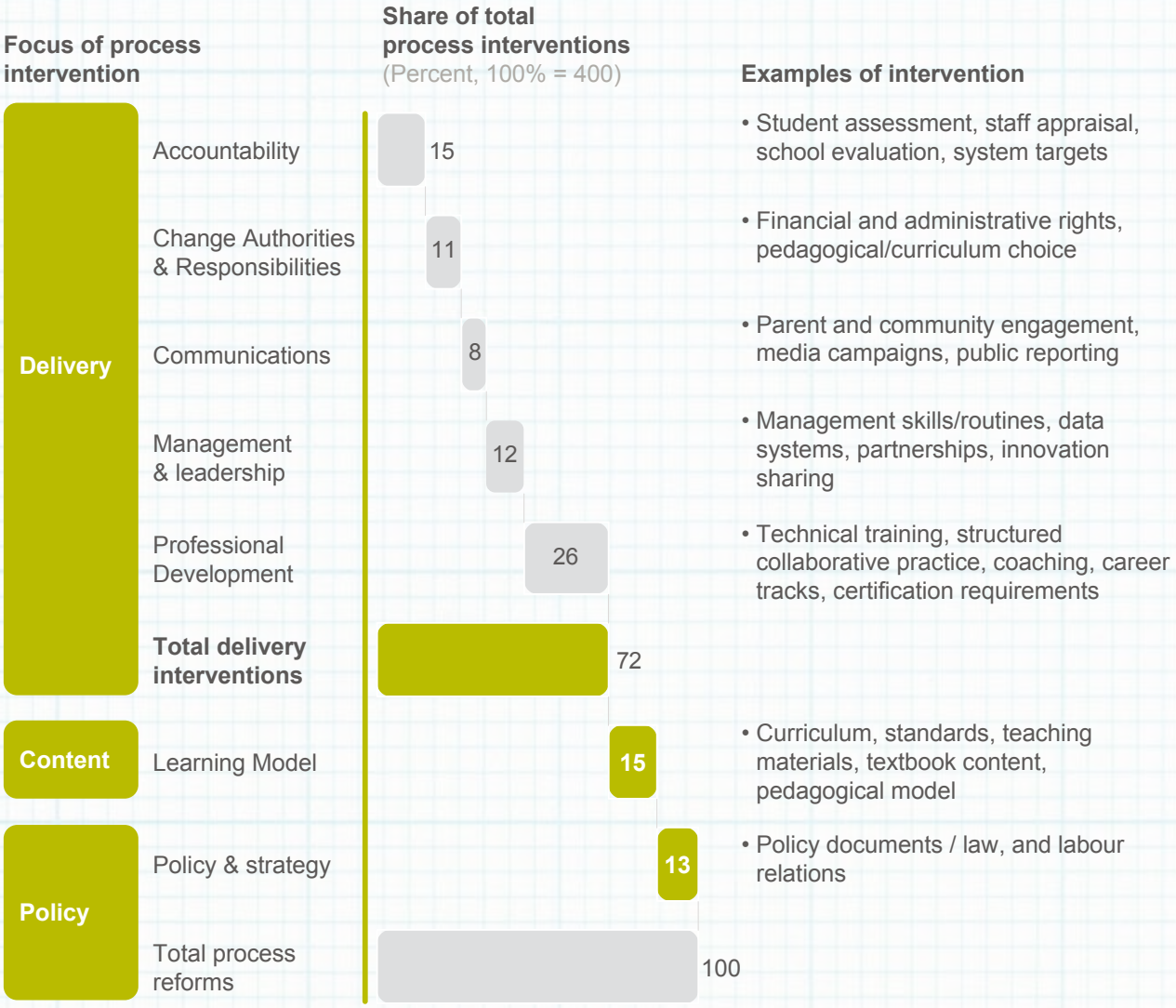
Source: McKinsey & Company interventions database

Remuneration adjustments are made at each improvement journey stage. However, the actual nature of these interventions varies considerably depending on the journey stage (Exhibit 23). Systems on the poor to fair journey give rewards to teachers or schools that meet proficiency targets. For example, Madhya Pradesh gave an additional month's salary to teachers who were able to achieve and sustain literacy targets for six months or more. Western Cape gave USD 2,000 to the highest-performing school in each wealth quintile in each of its eight school districts. In contrast, systems at the fair to good and good to great improvement journey stages sought to provide teachers with a respectable base salary (in relation to GDP per capita) and to tie the final salary to the particular point the individual teacher's position on their professional career path (see Chapter 3). As discussed above, these systems also sought to keep teachers motivated by improving their work environment and raising the level of esteem of the education profession. Lastly, systems on the great to excellent journey seek to ensure that teacher salaries significantly exceeded the national GDP per capita. These systems recruit top-performing students into the teaching profession, and so aim to provide competitive remuneration relative to other professions. For example, in Singapore, teachers currently receive a salary that is one notch higher than that of civil servants with equivalent qualifications performing general administrative functions. In addition, top-performing teachers receive a bonus that is the equivalent of up to three month's salary.[20]

Similarly, as with remuneration, technical skill building also manifests differently in the different journey stages. As systems progress, they shift from a small number of standardized "broadcast" training programs to a larger number of personalized training sessions. For example, systems undergoing the poor to fair journey cascade standardized literacy and numeracy training programs that focus on supporting teachers in learning the prescribed content and pedagogy, whereas systems undergoing the great to excellent journey seek to provide a greater number of professional development hours, but allow teachers flexibility in selecting the topics that are most relevant to their own development needs. For example, in Singapore, where teachers are encouraged to participate in one hundred hours of training each year, they are free to select the training modules that fulfill their own interests and needs.[21] As systems move further along the improvement journey stages, the topics of the training programs expand beyond merely providing technical knowledge to encompass areas such as professional practice, leadership and management, and even interpersonal skills. Moreover, teachers increasingly take responsibility for improving each other's instructional practice in their schools, particularly as they become more senior (see Chapter 3).

Likewise, the nature of assessment varies according to the journey stage. Overall, assessments provide data for evidence-informed policy during each performance journey and so enable the system ➔

# Exhibit 25:
# ~75% of process interventions undertaken by our sample systems dealt with school delivery

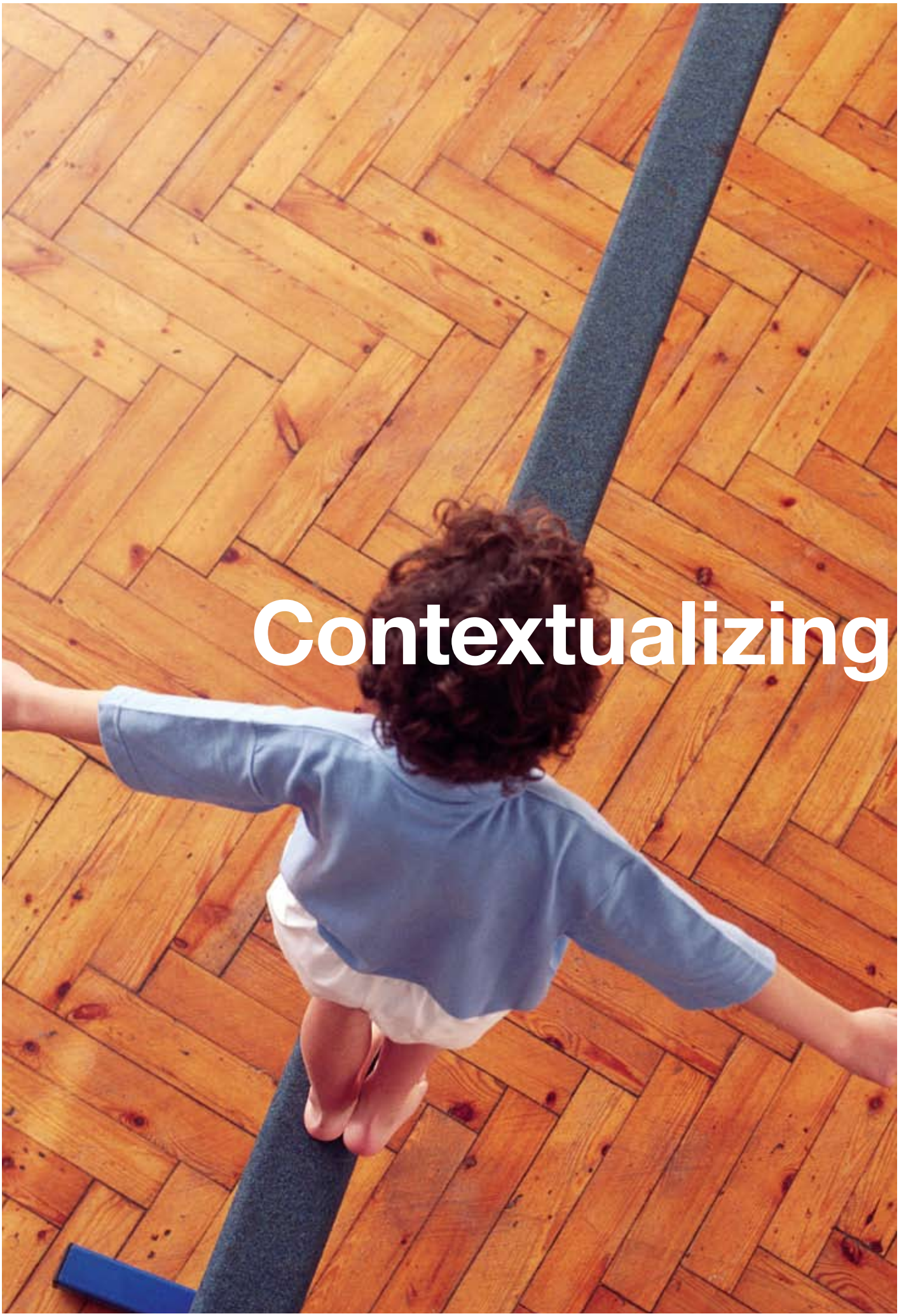



Source: McKinsey & Company interventions database

to continuously improve. In systems undergoing the poor to fair journey, assessment is carried out at the local or regional level and is narrowly focused on achieving specific targets for particular grades in the basics of literacy and numeracy. As the system progresses on its journey to the later stages, assessments broaden out to cover other subjects or additional grades. These systems also use assessments to conduct benchmarking research – for example, Lithuania developed TIMSS-like national assessments in science and math for grades four, six, eight, and ten in order to analyze trends of student learning. At the good to great and great to excellent improvement journey stages, the scope of such tests frequently expands to cover problem solving and more qualitative aspects.

One further observation spans all the various improvement journey stages; this relates to the categorization of interventions as process, structure, or resource-based. While each of our sample systems employed all three types of interventions at some point in their journey, we find that the vast majority of interventions are process-based (Exhibit 24). This might be explained in part by the fact that the menu of process interventions is greater than that for the other two types, or that several of the improving systems found that their contextual circumstances limited their ability to make resource and structure changes during the reform and therefore had to rely more heavily on process interventions. However, it also points to the fact that student outcomes can only improve with changes in classroom instruction. As a Boston system leader noted, "For student learning to improve, we had to improve teaching and learning practices in classrooms. And for that change to stick, the culture of classrooms and schools needed to change." In evidence of this point, only 15 percent of the total process interventions dealt with the content of system instruction (e.g. standards, curriculum), while ~75 percent dealt with the delivery of instruction. In terms of delivery, of this ~75 percent, the two most frequent interventions relate to professional development (~25 percent) and accountability (~15 percent), showing that enhancing these mechanisms is central to improving school system delivery at each stage of the journey (Exhibit 25). This finding is not to say that 'structure' and 'resource' reforms are unimportant; rather, it is clear from the reform journies described in this chapter that both have played shaping roles. We interpret the dominance of 'process' interventions as more a reflection of them being more used by improving systems.

This chapter has focused on how the improving systems at each particular journey stage select from a dominant intervention cluster, irrespective of culture, geography, and political context. In the following chapter, we now focus on the differences – looking at the different ways systems implement these interventions, adapting them to their own particular circumstances. ■



# Contextualizing

Improving systems tailor how they implement the intervention clusters in each performance stage to their context. The interventions are unlikely to achieve their full impact without this. Contextualizing tends to be aimed at gaining the support of stakeholder groups and, in particular, involves making decisions about when the system should mandate an action or when it should make compliance voluntary. To illustrate how the various improving systems take different approaches and use contrasting tactics in contextualizing implementation, we focus on the choices systems have made in three areas: professional development, the language of instruction, and student achievement targets.

## Break through, rather than break down

An important area of inquiry in our research has been to examine which aspects of school systems' improvement journeys are universally applicable and which are context-specific. First a definition of what we mean by context. Context has two main forms. The first type of context, which we addressed in Chapter 1, is current system performance (poor, fair, good, great, excellent) and its impact upon what a system does. The second type, which we address in this chapter, refers to the influence of history, culture, values, system structure, politics, etc. upon how the system implements the common interventions in their improvement journey.

We found that to implement the common cluster of interventions successfully, leaders use an array of strategies and tactics to accommodate the contextual realities in which they operate. We have chosen to focus this discussion on three particular aspects of contextualizing that occur across multiple systems and performance stages:

1 professional development requirements
2 language of instruction, and
3 student achievement targets

The most critical dimension differentiating the various implementation solutions the leaders of our sample systems used to address these aspects is in deciding whether to "mandate" or to "persuade," and in what combination. While mandating interventions delivers consistent quality across the system, persuasion enables stakeholders to feel ownership and autonomy. Managing this tension was a constant balancing act for the improving systems, the system's context being the fulcrum around which these decisions revolved.

## The guiding principles in mandating versus persuading

A characteristic of the systems we studied is the flexibility of their approach in achieving their reform aims: this flexibility is best demonstrated in their ability to oscillate between hard mandate and soft persuasion. Former Long Beach superintendent Carl Cohn posed two questions as he grappled with this style choice, "Reform in Long Beach was both bottom-up and top-down . . . The questions I always asked myself were how do I win over the people who do the work, and how do I also get them to feel urgency?"

Looking across our sample of 20 improving systems, a clear pattern emerges in the circumstances that determines when systems choose to mandate and when they choose to persuade. Mandating an intervention appears to be chosen as the dominant approach when at least one of the four following conditions occur in the system: 1) the desired change is considered "non-negotiable," an anchor point of the reform around which system leaders allow little or no compromise in execution; 2) there are few or no losers as a result of that particular change; 3) the credibility and stability of the system leadership and national government, and the historical and political context; and/or, 4) the pace of change needs to be rapid (due to political time pressure). In contrast, leaders tend towards persuasion when the inverse of the above is true (e.g. stakeholder groups are sharply divided, with clear winners and losers as the result of a particular change; or the leader or government presently has a tenuous hold on power).

Both these approaches have obvious benefits and drawbacks. Mandating enables fast action and fidelity of practice across the system, but risks stakeholder resistance. Persuasion allows stakeholders to gradually get used to the particular change and to feel real ownership over their decision, but risks complacency and the slowing of reform momentum. Our sample systems developed a set of tactics to mitigate the downsides of both approaches. When an intervention is mandated, the system leaders have gone to the front line to hear their views and to explain the change rationale. When persuasion is used, the system leaders work to build a critical mass to support for the change, while continuously reminding their stakeholders of the urgency of the desired change.

Across all systems, system leaders constantly sought to achieve the right balance between mandating and persuading. While they made ➔

great efforts to listen to stakeholders and embrace their concerns, they were equally clear in drawing the line when it came to pushing through a reform or in ensuring commonality across the system. As an Armenian system leader observed, "When I initially entered my job, I worried a lot about what different stakeholders think and I felt paralyzed. Then I visited other systems and I realized that stakeholders are not happy in any system. This is normal. I have come to understand that change creates resistance, no matter what. You cannot do my job if you are always worried about what other people think. You need to do the right thing for the system and have courage. Otherwise, you should not be in this job."

## Professional development

Several improving school systems have sought to raise the prestige of the teaching profession by increasing their professional development requirements and by creating clearly defined teaching career tracks; this is particularly the case in the fair to good and good to great performance stages. However, the implementation paths they have chosen vary markedly. We examine here three contrasting examples and the contexts in which they were applied. The first of these systems mandated a new teacher certification system, which significantly raised requirements, and implemented assessment of teaching practice; the second solely mandated the requirement that teachers should complete a certain number of professional development hours; the third system opted to make professional development completely voluntary.

Two contextual factors are important in explaining why Lithuania chose to mandate its new teacher certification system, including the assessment of teaching skills. First, during the early 1990s teaching was considered a low-prestige career choice in Lithuania – students placed teaching among the lowest options for university study, ranking it fifteenth. In the wake of the Soviet Union's dissolution, "Everyone wanted to be a trader … they made ten times as much money as a university professor," observes one senior Lithuanian school system leader. Second, the certification system was part of a broader move by the Lithuanian Ministry of Education to professionalize teaching. The ministry rolled out a new curriculum in 1996, including competency-based standards and instructional materials (new textbooks, teacher guides, and student workbooks). In the wake of this rollout, there was a strong imperative to ensure that teachers had the pedagogical competence necessary to deliver the curriculum, particularly in regard to cultivating application skills in their students. In parallel to mandating teacher certification, the ministry established a teaching career path and brought in commensurate salary increases for each successive rung on the ladder in order to increase the attractiveness of the teaching profession.

Following its pilot in 1996, Lithuania fully mandated teacher certification in 1998 – all teachers who started work before December 1, 1994 were required to obtain certification by December 31, 2001.[22] The Ministry of Education outlined four professional designations: classroom teacher, senior teacher (with responsibility for coaching other teachers in the school), "methodist" (with responsibility for coaching teachers across the district), and expert (with responsibility for coaching teachers at the national level and for supporting the writing of the national curriculum). Each step had explicit skill and tenure criteria and a salary increase of approximately ten percent accompanied each step. Teachers were certified through assessments conducted by their school, district, or a national committee, depending on the professional designation; three criteria were used for each subject taught: the teacher's instructional practice (based on observation), the fulfillment of in-service training requirements, and their knowledge of teaching theory. Recertification takes place every three years. By 2005-06, over 75 percent of Lithuania's 46,865 teachers were certified, with the vast majority of the remainder falling outside the stipulated tenure band.

In contrast, Hong Kong 'strongly recommended' that its teachers should comply with undertaking a stipulated number of professional development hours over a specified time period.[23] In 2000, the education system leadership focused its school system reforms on increasing school-based management skills and reducing performance variation across schools. Teacher and principal capacity building were an integral part of this effort; in 2002/03, Hong Kong recommended all its

teachers and principals to fulfill 150 professional development hours every three years.[24] Schools defined the mode and content of these professional development activities, as well as had autonomy in monitoring the implementation of teachers' professional development. "It was easily accepted by teachers and principals because we were just asking them to comply with something they were already doing and the requirements could be fulfilled in a flexible manner ... in fact, teachers regularly over deliver on their professional development requirements," says a Hong Kong senior educator. The flexibility in the system is evidenced, for example, by the option for teachers to fulfill professional development requirements with in-school development activities, such as department meetings, class visits, and attending/hosting lectures. Similarly, formal courses offered by the Education Bureau, teacher training institutions, or professional associations, could be credited to the teacher's professional development account, so long as the teacher was present for more than eighty percent of the course.

Hong Kong opted for an input-oriented certification process rather than a skills assessment because, in the words of one Hong Kong system leader, "We have a high diversity in Hong Kong schools due to the different curriculums on offer ... It is difficult for us to make a blanket imposition across all our schools. But we wanted teacher and principal professionalism to be clearly on the school radar screen and something that was in their dialogue on a regular basis." Since the 1960s, the vast majority of Hong Kong's schools have been publicly funded but privately run by charitable trusts and Christian missionary organizations, which have significant autonomy. "Our philosophy is that we cannot force schools to do something that they don't want to do, so we tend to use only support mechanisms, little pressure," says a system leader. This is one reason that when establishing its professional development requirement, the Education Bureau embarked on a campaign to appeal to teacher values, in order to encourage teacher support for the new requirement. Says a Hong Kong educator, "We said to teachers that if you consider yourself a professional, this is what you do." ➔

In 1999, Poland opted for a very different path, making professional development voluntary and at the discretion of each individual teacher. "It is very difficult to impose anything on anyone in Poland. We have a very strong ethos that the center should not tell people what to do ... this is a reaction to our centralized past with communism and martial law. We wanted teachers to feel freedom to participate in professional development, so we decided it should be voluntary, not obligatory ... our teachers would otherwise have resisted this change," says a Polish system leader. In addition to this historical legacy, stakeholders point to two other contextual reasons that drove the decision to make professional development voluntary rather than mandatory. Firstly, in the early 1990s, only the top ten percent of Polish students in their secondary school cohort could progress to university. Teacher candidates were drawn from this cohort, as they needed to have a Bachelor's degree in order to be able to teach. In the words of a Polish educator, "Twenty years ago, we recruited our teachers from the top ten percent of the graduating class, so the caliber of our teachers was already high." One stakeholder explains the second reason as follows: "The magnitude of the structural changes we were making to the system in 1999, by introducing an additional year of general education, was so huge that teachers themselves wanted help in order to cope with it."

Having decided that professional development should be voluntary, the Ministry of Education embarked on a two-pronged approach to encourage teachers to develop their skills. Firstly, the Ministry funded the creation of teacher development centers (TDC) across Poland. The centers were managed by Poland's sixteen regions and were responsible both for diagnosing the training needs of their local teacher populations and for offering in-service training courses accordingly; these courses were taught by TDC employees, professors from regional Colleges of Education, or high-performing regional teachers. The decision to place the governance of the TDCs in the hands of the regions was heavily influenced by the government-wide decentralization drive in 1999. Secondly, in a similar manner to Lithuania, Poland dangled a carrot; in 1999, it created a four-level teacher path with accompanying salary increases to motivate teachers to enroll in in-service training programs; their completion accelerated the teacher's progress in their career track. "We were concerned that once teachers became certified in the two higher levels and their job security was complete, they would stop progressing their skills, so we also established a bonus system which could reach up to twenty percent of total salary," says a Polish system leader. Bonuses are awarded to high-performing teachers by the school principals, subject to the guidelines articulated by the various regions.

Each of these three systems has had the same objective in seeking to increase the professionalism of their teachers. Each chose a very different blend of mandating and persuading in their implementation process, based on their particular context. Lithuania mandated both input and output because it viewed teacher certification as non-negotiable – this was necessary to accompany its curriculum changes. Moreover, its teachers were largely supportive because the scheme enabled them access to a higher role and salary, based on demonstrated competence. Hong Kong placed a "soft" input mandate to encourage further study, based on what its teachers were already doing, and so easily garnered their support. The structure of their school system made its leaders believe that a harder mandate would have been difficult to enforce, and might have resulted in broad stakeholder resistance. Poland chose not to mandate either input or output but to use persuasion. This was due to its concerns about potential teacher resistance to any other course due to political context; the system's leaders believed that this approach would be successful because teachers possessed a high base level of skill, the new career track created incentives to stimulate demand for training, and the ongoing curriculum changes were sufficiently large to ensure teachers would want to access in-service training support.

## Language of instruction

The language of instruction can be one of the most contentious issues in any school system because it taps into sentiment about personal and national identity. Language is an important issue in both Hong Kong and Singapore: each chose a different path to bring resolution in their respective contexts. Both grappled with these issues regarding

the language of instruction as they pursued interventions regarding literacy and pedagogical foundation. Hong Kong opted for a persuasive and flexible approach; Singapore chose top-down mandate. Both systems clearly recognize both the importance of English to the future employability of their students, and of the mother tongue to cultural cohesion and identity.

Hong Kong's Education Department outlined the importance of both Chinese and English to its student population: "Hong Kong is an international city as well as a business, financial and tourist centre … On the one hand, Chinese [Cantonese in the main] is the mother tongue and the language medium for everyday communication for the majority of our population. On the other hand, we recognize the importance of English proficiency in Hong Kong's continued development and prosperity. Hence, the aim of Government's language education policy is for our young people to be bi-literate [i.e. to master written Chinese and English] and trilingual [i.e. to speak fluent Cantonese, English and Putonghua]."[25] A series of public consultations sponsored by the Education Commission and dating back to 1984 had recommended Chinese as the medium of instruction, using the rationale that students learn most effectively in their mother tongue.

In the period leading to Hong Kong's transition in sovereignty in 1997, a broad stakeholder consultation process on the medium of instruction was undergone; the Education Department received 106 written submissions from school councils, educational bodies, school principal unions, teacher' associations, sponsoring bodies, educators, parents, student bodies, and other members of the public. The public debate was deeply divided, with teachers, who spoke little English, advocating for Chinese as the medium of instruction, while parents, who were concerned with the employability of their children, advocating that it should be English.

Following deliberations, the Education Department of the newly established Hong Kong government ultimately issued guidance in September 1997 that opted for persuasion, articulating that Chinese was the preferred medium of instruction but that schools were allowed to use English if they could justify doing so: "All local public sector secondary schools should, on the basis of the principles in the MOI [medium of instruction] guidance, examine their own conditions to determine the MOI appropriate to the needs and ability of their students. Starting with the Secondary 1 intake of the1998-99 school year, Chinese should be the basic MOI for all local public sector secondary schools. If a school should, after careful deliberation, intend to adopt English as MOI, the school must provide sufficient information and justification to Education Department to support such choice." Further areas of flexibility were also specified in the Education Department guidance:

At senior secondary levels, the MOI policy may be applied with more flexibility. Exceptionally, ➔

schools meeting requirements may, with Education Departments 's agreement, use English as MOI for some subjects.

a At sixth-form levels, schools may choose the MOI which best meets the needs of their students.

b For the subjects of religious studies, cultural, commercial and technical subjects, individual schools may choose the MOI which best meets their circumstances.

c After two rounds of applications and assessments, about a quarter of secondary schools (114 schools) were granted the right to use English as the medium of instruction; the other three quarters of secondary schools were required to adopt Chinese as the medium of instruction.[26]

At the same time, concern for teacher skills in English was high in wider Hong Kong society. Therefore, as part of the teacher professionalization drive in its Year 2000 reforms, the Education Department mandated certification for its English teachers, whereby all existing and new English teachers had to sit an exam, within a six-year period, which would qualify them to teach English. Against a backdrop of teacher concerns about job losses due to the shrinking Hong Kong student population, as well as fears that this certification process was just the first step towards making English the medium of instruction, the teacher unions protested strongly against the mandate.

For the first time in Hong Kong's history, teachers took to the streets in protest. At this point, the Hong Kong government introduced a new tactic, giving parents and employers a voice in the debate; both groups strongly advocated English certification. The Education Bureau stood firm in its mandate, and English teaching certification proceeded apace.

Singapore, in contrast, mandated its decision about the language of instruction top-down. Language has always been an important issue in Singapore. From the early days of Singapore's self-government in 1959, its schools taught in the four official languages (Chinese, Malay, Tamil, and English), all of which had equal status in order to facilitate its objective of building a cohesive, multi-racial society. Parents chose the language of instruction for their child's education; in addition, every student was required to learn a second language. In 1966, the government moved to mandate bilingualism; all students were required to study English, either as the first or second language, as well as their mother tongue, from primary school onwards. By 1969, 60 percent of all students were enrolled in English streams, 33 percent in Chinese, 6 percent in Malay, and 1 percent in Tamil.[27] At the time, the government viewed its 1966 bilingual policy as the cornerstone of Singapore's economic and social prosperity. Looking back at this decision, the then Prime Minister Lee Kuan Yew noted, "If we were monolingual in our mother tongues, we would not make a living. Becoming monolingual in English would have been a setback. We would have lost our cultural identity, that quiet confidence about ourselves and our place in the world."[28] Singapore's first Cabinet held a strong belief that the future of science and technology would be written in English and so its students therefore needed to be adept in the English language from an early age.

In 1978, the Goh report found that Singapore's bilingualism policy had not been "universally effective." Less than 40 percent of the student population had achieved the minimum competency level in two languages. The Goh report proposed streaming student by competency level from Primary 4 onwards; as part of this reform, the weaker students who struggled in learning two languages would focus on only one language (English), while students who excelled in two could choose a third language in addition to English and their mother tongue.

In 1983, the government mandate moved one step further, announcing that English would be the medium of instruction from 1987 onwards in all subjects, except the mother tongue, from Primary 1 onwards. This policy was consistent with the reality on the ground of declining student enrollment in Chinese, Tamil, and Malay schools, as the majority of parents saw English as offering the greatest employment opportunities for their children.

At the time, the magnitude of the challenge facing the school system in making this switch was immense, as few families spoke English among themselves: in 1990, just 18 percent of student

households spoke predominantly in English at home, while 65 percent spoke Chinese, 14 percent spoke Malay, and 3 percent spoke Tamil. Alongside the introduction of this mandate, the Ministry of Education also took steps to embrace its large, Chinese-speaking stakeholder group – teachers in Chinese schools were converted by the ministry into teachers of Chinese as a second language, while steps were taken simultaneously to increase the quality of instruction in Chinese.

Why did Hong Kong choose to use a degree of flexibility in implementing its medium of instruction policy while Singapore opted for mandate? Our interviewees highlighted four possible reasons.

First, school structure: Hong Kong has a privately operated, publicly funded system, whereby the majority of public schools are operated by private entities (charitable trusts and Church missionary societies).[29] As such, Hong Kong could only alter school and teacher behavior through the combination of regulation and incentives. In contrast, Singapore's Ministry of Education had full and direct control over its schools.

The second reason the interviewees cited is stakeholder alignment. Hong Kong experienced a range of contradictory and divisive views (in particular, between teachers and parents) in relation to the choice of the medium of instruction; in Singapore, stakeholder groups were more unified, and active steps were taken to appease the large, Chinese-speaking stakeholder group.

The third reason relates to what the two systems consider to be non-negotiable: in Singapore, student literacy, which was held to be non-negotiable, was being put at risk due to its bilingual policy, creating an urgency for mandate. Hong Kong also faced similar issues. Interestingly, English teacher certification was the one moment in its debate on the medium of instruction when Hong Kong chose to mandate rather than persuade, despite significant teacher protest. What made it veer from its usual path? Interviewees offer the following explanation: the professionalization of teachers was a hallmark of the 2000 reforms, making teacher competency non-negotiable. This was all the more the case for English skills, which affected Hong Kong's economic prosperity.

The fourth and final reason for the differences between the two systems was political: during the medium of instruction negotiations, the Hong Kong government was in transition (leading up to the transfer of power in 1997) and so had to tread gently, whereas Singapore's government has undergone a long period of continuity and was relatively invulnerable to criticism.

## Targets

All of the school systems in our research sample use data proactively and regularly to ensure that they understand which schools or students are falling behind or moving ahead, and to assess whether their reform efforts are having the intended effect upon students' learning outcomes. However, only a subset of our sample systems translate this data into quantitative targets at both school and classroom level. Even more striking is that this subset falls into just two groups: 1) systems of the U.S., England, and Canada (including Aspire, Boston, Long Beach, and Ontario); and, 2) systems making the journey from poor to fair (Madhya Pradesh and Minas Gerais). In these two groups, systems give prominence to the agreed targets and share them widely with stakeholders, and in most cases, with the broader public. In contrast, the Asian and Eastern European systems in our sample refrain from setting quantitative targets, preferring to share performance data with individual schools, engaging them in a private dialogue about how they can improve. In this latter group, only system-level data is made available publicly.

All of the improving systems in our sample had high performance expectations to schools. Why is it that the Anglo-American systems and those on the poor to fair journey chose to express these expectations as achievement targets for schools and students, whereas the Asian and Eastern European systems opted to avoid targets and persuade schools to improve behind the scenes? There can be no simple answer to this question, other than to say that is the outcome of their "socioeconomic, political, and cultural context." Culture, history, and system values intertwine, resulting in system leaders choosing different tactics in how ➔

they use data. For example, several U.S. interviews noted that the "No Child Left Behind Act" of 2001 (which requires states to set student standards and conduct assessments of basic student skills in order for them to receive federal funding for their schools) has played a significant role in fuelling the target culture of the U.S. In contrast, as one Singaporean leader noted, "Singapore's value system is that first priority is the nation, second is the organization, and third is the individual. We do not need to use targets because the greater good of the nation motivates our people to always work towards improvement." This leader went on to further speculate that in systems where the reverse value system is the case, where the individual ranks first, targets might be an important way to create common ground. Again, the context in Eastern Europe is very different, as one Eastern European leader noted, "Our centralized past makes it very difficult to set top-down targets on anything ... Our schools and educators want freedom to do what they think is best. Performance targets would be viewed as excessive state control and would be rejected."

The Anglo-American and poor to fair systems follow a consistent pattern in announcing ambitious student achievement targets publicly, conducting regular testing to assess performance against these targets, and then making assessment results visible to stimulate school motivation and discussion on how to improve. "We are here for the kids. This is not about protecting the adults in the building," says a U.S. system leader. Aspire, for example, sets an 85 percent target for each of its class learning objectives and conducts tests every two to three weeks to assess progress. These results are then posted on the classroom walls for all its students, parents, fellow teachers, and staff to see. This transparency results in regular discussions about why the performance varies across classes in the same school, and across schools. A North American system leader says, "There is no shame in transparency; the only shame is in not asking the question of why a classroom is failing to deliver ... If a teacher gets a classroom average of forty and is trying to improve, then we will support him or her. If it persists, then that teacher does not belong here. Our kids come in with so many strikes against them, we don't have time to waste." Across the Atlantic, achievement targets were used in England in the same spirit, to focus schools on delivering literacy and numeracy gains for their students. It also published league tables on an ongoing basis to rank school performance on assessments and inspections; repeated poor performance could result in the schools being shut down.

There is some variation, however, in the extent to which the agreed targets are made public in these systems. While Aspire, England, and Long Beach schools all made school and classroom data and targets public, Boston established annual student achievement targets, and held principals accountable for delivering them – but refrained from announcing these targets publicly.

Systems on the poor to fair performance journey – Madhya Pradesh, Minas Gerais, and Western Cape – also turned to making quantitative performance data public as an anchor of their improvement journeys. For example, in Minas Gerais, the Department of Education set a literacy improvement target for the system, and then cascaded it down to each school based on its performance in a state-wide literacy test. The achievement targets were first agreed with school principals, after which each school became formally committed to meeting their targets for the coming three years by signing a "Target Terms of Agreement." The targets, and the school's performance in relation to them, were communicated widely to the public. Similarly, Western Cape engaged in public discussion about school performance data with the districts, schools, and the community during its annual performance road trip through the province.

Leaders across these three systems engaged in poor to fair journeys noted that targets helped them to align stakeholders on a small set of priorities, and gave them a simple metric regarding their progress in making improvements that could be easily communicated and understood by the public. A similar sentiment was echoed by a U.S. system leader: "When there is time pressure to deliver results, and the stakeholder environment /is fractious, sometimes agreeing a number is the best way to get people focused."

The Asian and Eastern European school systems are at the opposite end of the spectrum to the Anglo-

American ones. "We have never used targets... No good for our students could ever come from making school data public and embarrassing our educators," says an Asian system leader. Across the board, the Asian system leaders expressed a two-fold rationale for why they do not set targets with schools, and why they only make system-level data public.

The first point the Asian leaders voiced was a belief that top-down target-setting is at odds with developing holistic school capacity; they view this as a trade-off between their desired focus on processes (e.g. school excellence, teaching and learning practices) and that on targets. "We want our schools to focus on getting the process right. If they follow the process, they will get good results. But if they focus on targets, they can end up taking shortcuts in the process." Even during the period when the Asian systems in our sample were on the poor to fair journey, none set quantitative targets for their schools.

The second reason articulated by the system leaders as to why the Asian systems do not make performance data public, is that they believe that "naming and shaming" hurts both educators and system learning. "Making results public demotivates staff and results in their paralysis ... they stop being open to learning and trying new things. Instead, they would focus on protecting themselves and finding ways to make their students look good on tests."

The Eastern European system leaders, in coming to terms with the past centralization of the Soviet system, have a very different rationale for reaching the same conclusion. An Eastern European leader says of their approach, "Sustainability of improvement lies in the professionalism of teachers and principals. We cannot force them. We have to convince them to want to improve if anything is going to change in our schools."

Both the Asian and Eastern European systems in coming to similar conclusions, despite their very different contexts, opt to share school assessment and/or inspection data privately with the school, rather than broadcasting it. Usually this takes the form of indicating a school's relative position in relation to other schools (e.g. school X is placed fifth on the national grade six math assessment). Some systems also reveal a school's "statistical neighbors," whereby the performance of schools or districts with similar student demographics are compared anonymously, thereby taking away the excuse that "my students are different." The sharing of data in this manner serves as the foundation for a dialogue between the school and the ministry on how the school will set about improving its performance, and is accompanied by a menu of different types of support that the ministry can offer schools in their improvement journey.

While the first major decision of a system leader is in deciding what to do to raise student outcomes (which interventions to make), the second is in deciding how to implement them (contextualizing the interventions). The three examples discussed here – professional development, language of instruction, and targets – illustrate the broad diversity in approach and tactics that systems employ to implement the same interventions. In particular, the spectrum of choice along the "mandate or persuade" continuum has been central to how system leaders ensure stakeholder's support for the reforms. The leaders of these improving systems show that they are highly attuned to their system's context, working with this and around it, contextualizing their interventions to achieve enhanced performance. ■

# Sustaining

For a system's improvement journey to be sustained over the long term, the improvements have to be integrated into the very fabric of the system pedagogy.

We have identified three ways that improving systems do this: by establishing collaborative practices, by developing a mediating layer between the schools and the center, and by architecting tomorrow's leadership. Each of these aspects of sustaining improvement is an interconnected and integral part of the system pedagogy.

Lee S. Shulman articulated why pedagogy is so fundamental:

*"Signature pedagogies are both pervasive and routine, cutting across topics and courses, programs and institutions... Pedagogies that bridge theory and practice are never simple. They entail highly complex performances of observation and analysis, reading and interpretation, question and answer, conjecture and refutation, proposal and response, problem and hypothesis, query and evidence, individual invention and collective deliberation... One thing is clear: signature pedagogies make a difference. They form habits of the mind, habits of the heart, and habits of the hand."*

This chapter reviews our findings about how improving school systems successfully embed the improvements they have made to ensure that the improvement can be sustained over the longer term. It is clear from talking to the leaders of the 20 systems studied here, and more widely, that sustaining change requires altering the very fabric of the system – changing not just the way teachers' teach and the content of what they teach but how they think about teaching. Sustaining improvements focus on producing a new professional pedagogy.

There are three important approaches improving systems introduce for sustaining their new pedagogy: collaborative practices, a mediating layer, and the architecting of tomorrow's leadership. We explain these in reference to a personal computer (PC), with which all of us have at least a passing familiarity: like a school system, all parts of a properly functioning PC have complementary roles in producing the desired outcomes.

- Collaborative practices: the user interface. A user interface is one of the most important parts of any PC system because it determines how easily you can make the computer do what you want. A powerful program with a poorly designed user interface is of little value. Collaborative practices embed routines of instructional and leadership excellence in the teaching community, making classroom practice public, and develop teachers into coaches of their peers. These practices are, in turn, supported by an infrastructure of professional career pathways that not only enable teachers to chart their individual development course but also help to share their pedagogic skills throughout the system. Collaborative practices shift the drive for improvement away from the center to the front lines of schools, helping to make it self-sustaining.

- The mediating layer: the operating system. In a PC, the operating system is the most important program – the one that runs all others. Operating systems perform basic tasks, such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk, and controlling peripheral devices such as disk drives and printers. It is tied closely to both the user interface and to the central processor and mediates between the two. The mediating layer in the school system performs a very similar role. As the school systems we studied have progressed on their improvement journey, they have increasingly come to rely upon this mediating layer between the center and the schools for sustaining improvement. The mediating layer gives three things of importance to the system: it provides targeted hands-on support for schools, it acts as a communications buffer between the school and the center, and shares and integrates improvements across schools. ➔

- Architecting tomorrow: the central processor. The central processor carries out all the calculations a PC undertakes. It reads and executes program functions and decides what information goes where within the PC. It is the central processor that ensures that what happens next time is the same thing that happened the last time. In a school system, the continuity of the system's leadership plays a similarly important role because the priorities, drive, mindset and resourcing of change are highly influenced by its leaders. Sustaining system improvement, therefore, somehow needs to traverse smoothly from one leader to the next, so that change becomes evolutionary in nature. We observe that the most successful examples of continuity come from systems that are always architecting tomorrow's leadership today.

We will now look at examples of how each of these elements in a system's pedagogy translates into practice.

## Collaborative practice: The user interface

Collaborative practice is all about teachers and school leaders working together to develop effective instructional practices, studying what works well in classroom, and doing so both with rigorous attention to detail and with a commitment to improving not only one's own practice but that of others. Collaborative practice is the method by which a school system "hardwires" the values and beliefs implicit in its system into a form manifest in day-to-day teaching practice.

Systems that have embedded these practices reinforce them through publicly acknowledging their teachers' proficiency and expertise levels within the system career track – a teacher's promotion carrying with it not just the recognition of their knowledge but of their compliance with the right pedagogical values – as well as the responsibility for sharing this expertise with others. We look first at peer-to-peer collaboration and then at career pathways.

Michael Fullan explains the power of collaborative practice, what he terms "collective capacity" in this manner:

*"Collective capacity is when groups get better – school cultures, district cultures and government cultures. The big collective capacity and the one that ultimately counts is when they get better conjointly – collective, collaborative capacity, if you like. Collective capacity generates the emotional commitment and the technical expertise that no amount of individual capacity working alone can come close to matching...*

*The power of collective capacity is that it enables ordinary people to accomplish extraordinary things – for two reasons. One is that knowledge about effective practice becomes more widely available and accessible on a daily basis. The second reason is more powerful still – working together generates commitment. Moral purpose, when it stares you in the face through students and your peers working together to make lives and society better, is palpable, indeed virtually irresistible. The collective motivational well seems bottomless. The speed of effective change increases exponentially ..."*[32]

## Peer-to-peer professional collaboration

Collaborative practice is about teachers and school leaders working together to develop effective instructional practices, studying what actually works in classroom, and doing so with rigorous attention to detail and with a commitment to not only improving one's own practice but that of others as well. As Lee Shulman notes, "A feature of signature pedagogies is that they nearly always entail public performance."[30] In his synthesis of over 50,000 studies and 800 meta-analyses of student achievement, John Hattie drew one major conclusion: "The remarkable feature of the evidence is that the biggest effects on student learning occur when teachers become learners of their own teaching."[31] This is the essence of collaborative practice: teachers jointly engaged in an empirical, routine, and applied study of their own profession.

A remarkable effect of collaborative practice is that it serves as a mechanism of peer accountability, substituting for other formal accountability measures such as teacher appraisals or requalification (Exhibit 26). Upon embarking on our study of improved school systems we had anticipated finding that systems implement interventions in teacher support and accountability in equal measure. The picture that emerged from across these 20 systems was very different. Teachers were overwhelmingly the greatest recipients of support, being the direct recipients of 56 percent of all support initiatives (e.g. professional development and coaching), but were the recipients of just three percent of accountability measures (i.e. teacher appraisal or proficiency assessments). How then did systems hold teachers accountable, if not through instituting appraisals? The answer to that question lies in two parts: the first is that teachers across these systems were held accountable through the learning of their students. The focus of these systems was on what students learned, not on what teachers taught, and that is reflected in the fact that student assessments represented 44 percent of accountability measures but direct teacher appraisal represented just three percent. However, even if student assessments were considered a mechanism of teacher accountability and these percentages are summed together, teachers would still receive a significantly higher share of support (56 percent) than that for accountability (47 percent). The second source of accountability was less formal but more powerful, and came from peers through collaborative practice. By developing a shared concept of what good practice looks like, and basing it on a fact-based inquiry into what works best to help students learn, teachers hold each other accountable to adhering to those accepted practices.

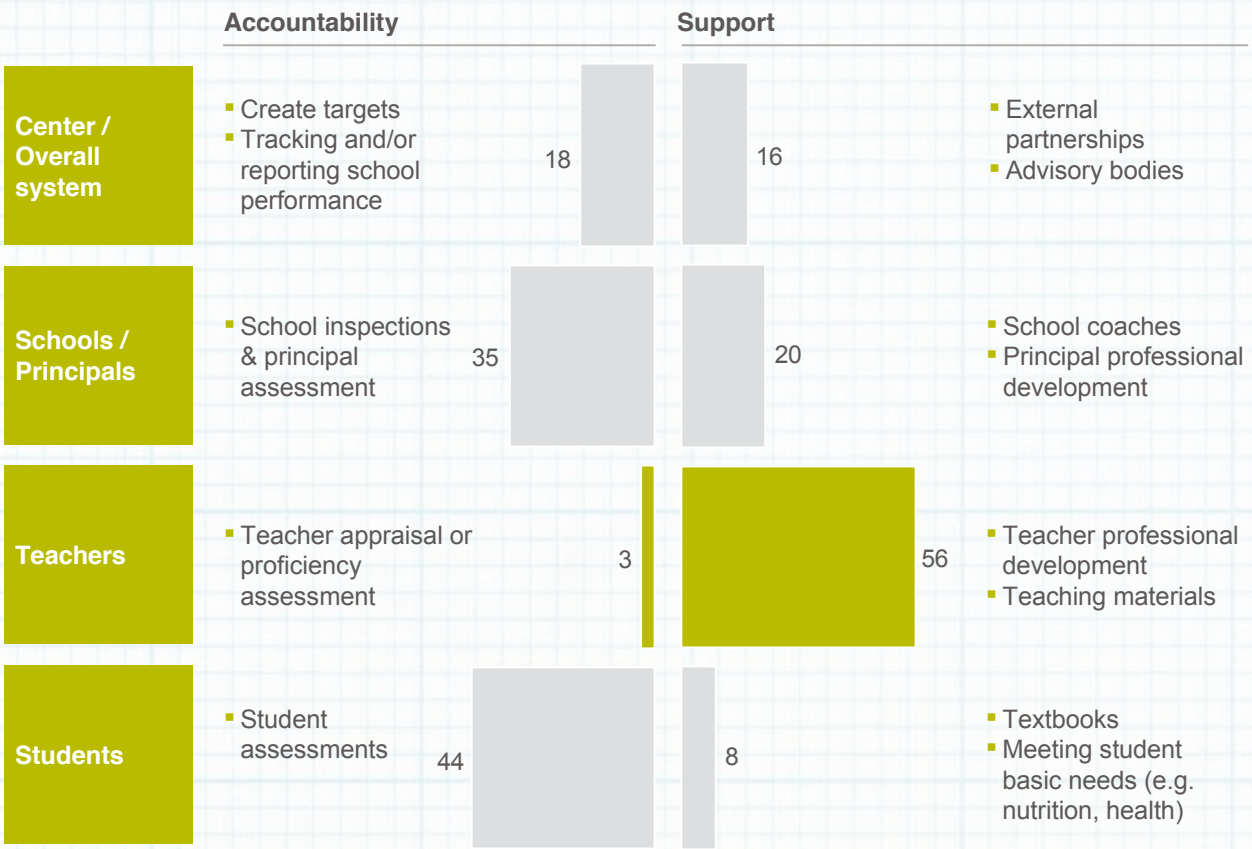
An account of how this can visibly manifest pedagogy – in our analogy, creating the "user interface" – comes from an educator in Ontario, who shares an account illustrating how peer accountability emerges from collaborative practice.[33] This is the story of a teacher who joined a primary school that had established the routines of collaborative practice as part of Ontario's literacy and numeracy strategy – these were professional learning communities through which teachers jointly reviewed student learning and developed teaching methods. In that teacher's first week in the new school two of his colleagues visited him and suggested he should use word walls because they had both found them to be effective. When, two weeks later, he had not yet put up the word walls, his colleagues visited him again, this time urging him more strongly to put up word walls, sitting him down to share why this was the practice in their school and the difference it had made for students. A few weeks later, by then well into the school term, the new teacher had still not put up his word walls. His colleagues stopped by again after school, this time simply saying, "We are here to put up your word walls with you and we can help you to plan how to use them." As professionals in that school, they had developed a model of instruction that they found effective and which had become hardwired as part of their values (a pedagogy), so they expected others to use it too. Their commitment was to all students and to their professional norms – not just to their own students in their own classrooms – and they were willing to hold each other accountable for using practices that they found effective. Together, the three of them put up the word walls.

We turn now to look at examples of how various systems have instilled collaborative practices in their schools.

Aspire Public Schools (APS), a set of charter schools in California, consistently and dramatically outperforms other schools in the districts in ➔

## Exhibit 26:
## Teachers directly receive 56% of all support interventions and only 3% of all accountability interventions

% of interventions in reform area directed at agent among all improved systems[1]




1 Accountability: N = 101; Support: N = 180

Source: McKinsey & Company interventions database

which it operates. Aspire is not selective about student intake and its schools have the same student profile as other schools in the districts. To take the example of two districts in which Aspire operates, Oakland and Stockton: whereas the district-wide average scores in California State API exams was 695 for Oakland and 694 for Stockton in 2008, Aspire students scored 775 and 833 respectively. Its students also tend to improve faster than other students, with API gains three to five times the average in 2007-08. As one Aspire leader reports: "We have moved through three performance horizons: first was our secret sauce of consistency in practice and accountability and that raised achievement; then we figured out that we should get targeted, with a very clear understanding of individuals student's needs and a plan to meet them; and finally we discovered the magic of effective joint lesson planning. We had been stuck below the 800-point, but once we started joint lesson planning, several of our schools passed through the 800-mark barrier. It was a remarkable breakthrough."

At the heart of Aspire's implicit values is a rigorous attention to data-driven improvement. The system has an almost religious commitment to empirically analysis of what works in practice and then applying it. At Aspire practices do not have to be perfect from the start, but they have to improve based on what works. Every conversation about what works is data based; without data, opinions lack credibility within the Aspire system. However, once armed with data, nothing is sacred. For example, in its early days Aspire adopted multi-age classrooms as a key component of its system, based on research that highlighted their benefits. When it discovered that the schools with multi-age classrooms were not showing improvement, it changed the model, keeping those elements that worked, such as the idea that classes should retain the same teacher for more than one year, while discarding the rest. Following these changes, performance in the schools rose sharply.

Joint lesson planning has become a cornerstone of Aspire's collaborative practice. Half of each Friday is set aside for teachers to work together in planning their lessons, during which time they review student progress and develop lessons to meet their students' needs. Teachers do this together, either in grade-based groups in primary schools, or in subject based teams in secondary schools, and wherever possible using coaches. There is a model for what makes a good lesson plan (part of Aspire's "operating system"; see below) that has been developed through experience. Lesson plans are expected to have eight specific elements and to be linked into the course and larger theme. Each plan is evaluated according to a rubric that explicitly defines what is expected of new teachers, basic teachers, proficient teachers, and distinguished teachers.

As visitors to a number of Aspire schools, the similarity between different schools and classrooms in their teaching practice was very apparent. This non-mandated uniformity is a product of collaborative practice. Instructional materials and methods are co-developed by teachers, tested in classes, and the results studied. What works well is shared widely and adopted by peers. What does not work is discarded. The expectation of teachers is not only that they should develop and employ effective practices in the classroom, but that they should share them throughout the whole system. Best practice therefore quickly becomes standard practice, adding to the pedagogy.

Aspire applies the same approach in relation to its instructional coaches. Early on, a great variety of coaching methods were used in its schools and there was not a lot of clarity about what worked best, so it collected feedback from its schools on what was helping students. From this data Aspire created standard guidelines for instructional coaching it calls the "the four acts of coaching." Coaches are expected to spend 85 percent of their time with schools and are prescribed a minimum number of touch points with teachers. They spend time with teachers in lesson planning and in "real-time teaching" – where the coach sits in the classroom when a lesson is underway, the coach wearing a microphone and the teacher a headset; the coach gives real-time feedback and guidance to the teacher while they teach to improve instruction. Collaborative practice is a feature of systems at the good performance stage, irrespective of culture. We encountered collaborative practice wherever there are high-performing schools. In Hong Kong, for instance, 50 hours out of the ➔

150 hours of the professional development required of teachers over three years is expected to be in collaborative practice formats, including study circles and mentoring. Hong Kong has also set up a School Support Network, which connects schools in working together to implement curriculum reform. In addition, it has formed professional learning communities among teachers to develop and disseminate effective instructional practices between schools. In Latvia, collaborative practice takes a slightly different form. One leading school has set up a "pedagogical lab" in which teachers develop lessons, record them on video, share them with peers, and discuss them. Each teacher is required to develop and share a minimum of three to four demonstration lessons per year. Teachers from other schools are invited to participate because the commitment to improving teaching practice is understood as a general commitment to the profession, not just to the school. In another example of collaborative practice, the Boston Public School district has introduced a "common planning time," in which teachers teaching the same subject and/or grade work together on lesson planning. Lastly, in England in 1999, the Department for Education and Skills appointed Leading Maths Teachers to deliver demonstrations of daily math lessons to primary school math teachers, modelling effective practice. Ten years later, in 2009, the concept was being extended to secondary schools.

The examples of collaborative practice we encountered were in systems on the journey from good performance onwards. At lower performance stages there is a heavier reliance on cascaded training for disseminating the basic pedagogy. Effective collaborative practice depends on teachers with strong capabilities. One Singaporean educator was clear on this point: "We could not have implemented professional learning communities as effectively in the 1980s. We did not have the skill levels in school for it and it may have backfired. However, our teachers and leaders are highly skilled now, and therefore we have shifted to peer collaboration more … and it works." While there may be cause to introduce professional collaboration irrespective of skill and performance level, in low-skill low-performance systems the experience of the poor to fair systems shows that the quickest way to produce a step-change in performance is to provide central guidance for instruction through the supply of scripted materials and cascaded training. However, as teacher capabilities rise, the distance between teacher and coaches in terms of their expertise levels reduces, ultimately making the teachers themselves the instructional experts in the system.

## Pathways for professional growth

Once a teacher has adopted the right approach – the system's pedagogical values – and has learned to manifest these in effective teaching practice, they become an invaluable asset to the school system, which then often seeks to embed this expertise by promoting such teachers to new roles. As teachers progress along the professional path, they assume responsibilities as educators, mentoring and leading other teachers, as well as in developing new curricula of the system. We now look at examples of how systems have achieved this.

Starting in 1996, Lithuania instituted a new teacher qualification system intended to professionalize its teachers while compensating them based on their skill level.[34] It had five qualification levels, giving teachers greater responsibilities at each step:

- Junior teacher: the starting grade for a teacher in an apprentice role.
- Teacher: teachers are designated as fully qualified for the role once they have undergone one year in the classroom; this designation is awarded by their principal of their school.
- Senior teacher: a teacher with two years experience: they are expected to coach other teachers in their school; their designation is decided jointly by their principal and the municipality.
- Methodist: a teacher who has been a senior teacher for five years and who has coached other teachers district-wide; their designation is decided by the municipality.
- Expert: a teacher who has been a "methodist" for seven years (and who therefore has been a teacher for at least 15 years), and who has coached teachers at the national level and contributed to writing the curriculum; their designation is decided by the Lithuanian

Teacher Qualifications Institute, a national council under the Ministry of Education, after nomination by their principal and endorsement by the municipality.

The assessment of teachers was initially carried out using demonstration classes (which were sometimes filmed); in addition, the teachers were required to undertake qualifying training courses and exams. Lithuania's teachers responded well, and student outcomes rose nationally. However, in the early years of this century the government enforced a public sector salary freeze and decentralized the methodist and expert designations to the school level, leading to some decay in the rigor of the qualification system. For example, while only 12 percent of teachers were methodists in 1998, this number rose to 20 percent by 2005.

The collaborative practices described here, supported by a system of professional development, can unleash sustained improvement; over time shifting the source of a system's improvement away from central leadership to the educators themselves. Teachers are in a position to sustain improvement because they draw motivation from seeing the impact on their own work, as well as from their ownership in shaping educational practice.

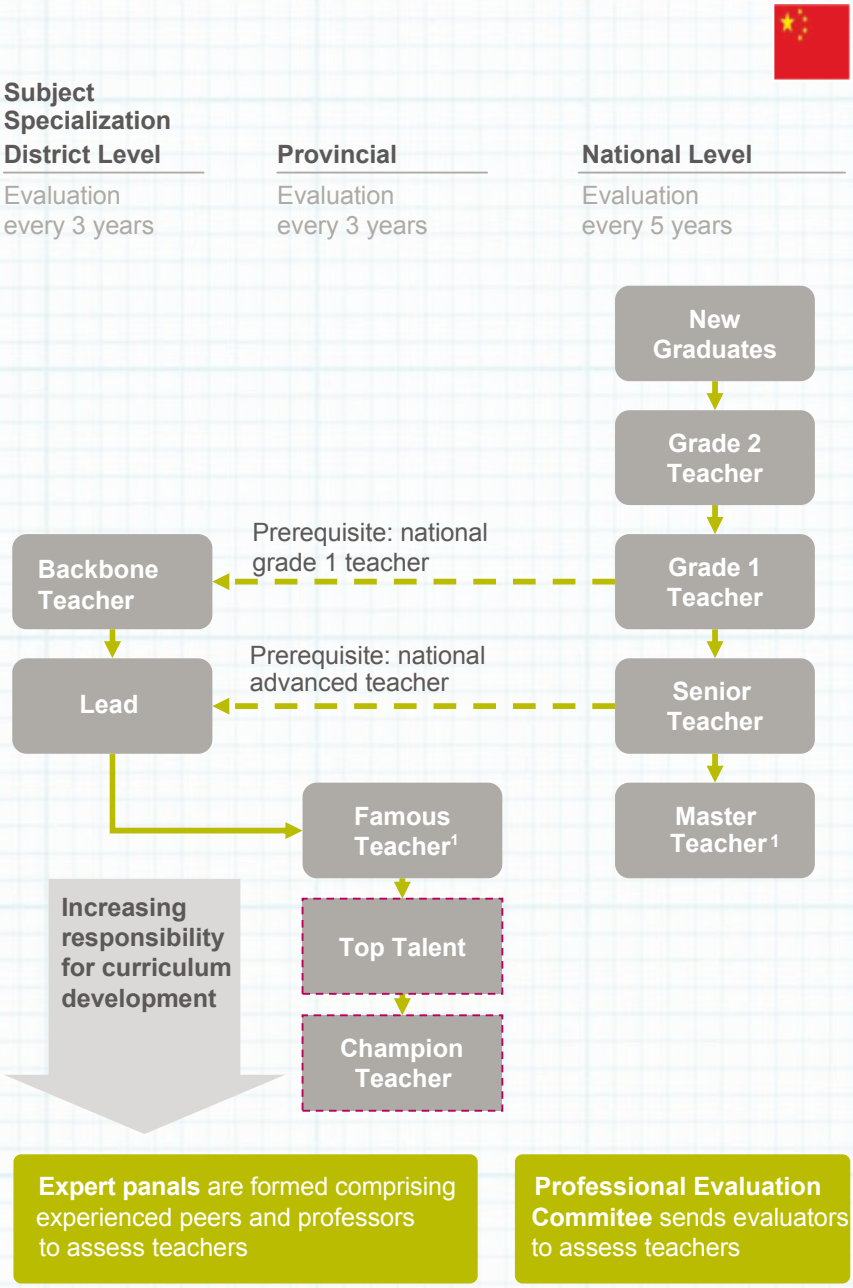
In a similar manner to how the language of a computer user interface determines the "look and feel" of a computer, the nature of a system's collaborative practice will determine its pedagogy. In the systems we encountered that had established strong routines of collaborative practice, system leaders bore witness to three changes that collaborative practice had brought about. First, it had moved their schools from a situation in which teachers were like private emperors, to one where teaching practice is made public and the entire teaching profession shares responsibility for student learning. Second, they report a cultural shift, moving from an emphasis on what teachers teach to one on what students learn. This shift results ➔

## CHINA

**In China, teaching and development teams, or *JiaoYanZu*, work together within schools and across schools to plan how the curriculum will be taught, to share learnings, and observe each other's practice. These teams serve as the pedagogical backbone of the school system. For example within schools in Shanghai, subject-specific teams (e.g. all third-grade maths teachers) meet each week to reflect on the past week and plan lessons for the following week. Subject group leaders (e.g. for history, math, science) also meet weekly to discuss how learning themes can be reinforced across subjects. In addition, subject-specific team leaders across schools in each district are required by the District Education Bureau to regularly visit other schools within the district in order to observe demonstration classes and share learnings. The objective of the JiaoYanZu is to cultivate shared ownership of teaching practices, to create consistency and to hardwire improved practices across the system.**

**China complements the *JiaoYanZu* approach with an integrated teacher designation and recognition scheme at the national, provincial (or municipal) and district levels (Exhibit 27). Each designation or grade has different requirements on the balance of teaching and development activities or responsibilities. The evaluation process for designation is carried out regularly by assigned evaluators and expert panels of experienced peers or professors. In Shanghai for example, the national designations run parallel to municipal and district designations, such that a teacher can be a "Grade 1 Teacher" according to her national designation, and a "Backbone Teacher" according to her district designation, with the former being a pre-requisite for the latter. The municipal level designation of "Famous Teacher" or MingShi, requires achieving the "Senior Teacher" status on the national scheme, and represents a critical role in the apprenticeship system of the profession. Each *MingShi* is expected to mentor a number of other teachers, meeting them each fortnight. MingShi are also provided with the resources to maintain offices at their school and to lead "development workshops" for curriculum development and professional development programs for teachers from across their district.**

# Exhibit 27:
# Teacher professional pathways in China




- Separate teacher designations at district municipal and national national levels
- National level designations are linked to subject specialization at district and municipal levels
  - E.g. Achieving Grade1 Teacher designation is a pre-requisite for recognition as a level Backbone Teacher
- Increasing recognition of status brings greater responsibility for mentoring and curriculum development

1 These teachers maintain a Workshop at their schools for curriculum development, teacher development and mentoring

Source: Interviews

both from an emphasis on studying student learning progress and from working together to develop approaches to improve that learning. System leaders believe this shift is absolutely necessary if all students are to learn successfully. Third, they report collaborative practice develops a normative model of "good instruction" – the pedagogy of the user interface – and makes teachers the custodians of that model. This is the characteristic of a true profession, in the same spirit in which doctors, lawyers, and accountants establish norms of good practice and hold each other accountable for following them. Lee S. Shulman describes this phenomenon in the following way: "Any signature pedagogy ... has a deep structure, a set of assumptions about how best to impart a certain body of knowledge and know-how. And it has an implicit structure, a moral dimension that comprises a set of beliefs about professional attitudes, values, and dispositions."

## The Mediating Layer: The Operaing Program

In formulating our research hypotheses, we anticipated that progress for the whole system would require improvements in both schools and at the center (i.e. in the ministry or head office). Student learning would not progress without improving what happened in classrooms, and whole systems of schools could not improve systematically and sustainably without changes in the support and stewardship provided by the center. What we did not anticipate, and what was not raised in the many preparatory discussions we had with educators before visiting the systems in this study, was the critical role that the mediating layer plays between school delivery and the center. In terms of our computer analogy, this role is akin to that of the operating system acting as a conduit and interpreter between the user interface and the central processing unit. We found that sustaining system improvement in the longer term requires integration and intermediation across each level of the system, from the classroom to the superintendent or minister's office (Exhibit 28). The operating system of the mediating layer acts as the integrator and mediator between the classrooms and the center.

This is not to suggest that school reforms should begin here. In every system we looked at, the first focus of school reforms was on the schools and the center. Efforts to strengthen the mediating layer usually came later, as the need for an active intermediary in delivering the system improvements became clearer. As one education leader in England reflects, "What happens in schools and classrooms has to always be the focus, and for us it was. However, in retrospect we could have recognized earlier how important the local education authorities were to improving what happened in schools and classrooms across the system. Once we figured that out, it made a big difference."

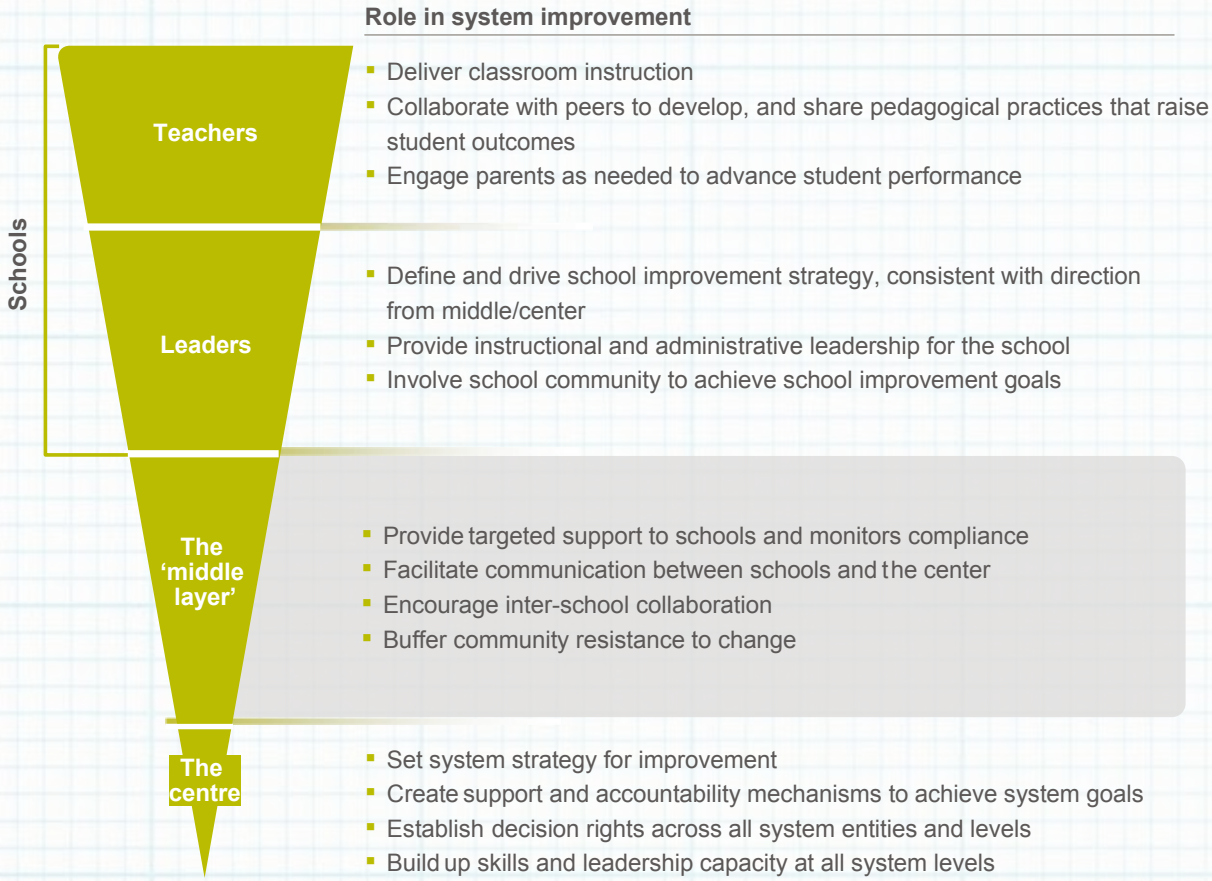
In several systems where the mediating layer already existed, its role in delivering improvement was strengthened: as was the case, for instance, in the local education authorities in England, the municipalities in Poland, the school boards in Ontario, the districts in the Western Cape, the regional and school-based support services to schools in Hong Kong, and the provincial offices in South Korea. In other systems where there was no such intermediary, such as in Singapore and Boston, a mediating layer (school clusters) was created afresh to meet the need for strengthening coordination and support across schools (Exhibit 28).

Each mediating layer has a common purpose, like computer operating systems, in interpreting, standardizing, and communicating. And like computer operating systems, there can be more than one approach to how this is done. We encountered four types of mediating layers among our 20 improving systems: a geographic mediating layer, school clusters, subject-based mediating layers, and level-based mediating layers.

A large proportion of the systems we studied had a geographically defined mediating layer designed to cascade administrative, financial, and instructional support for schools from the national/state/provincial level to the district/municipality level and, in the case of some the largest systems, one level beyond.

Other forms of the mediating layer were context-specific, responding to the needs and practices ➔

# Exhibit 28:
# System improvement requires intergration and coordination across every level



of the particular system. School clusters, such as those in Singapore and Boston, and subject-based mediating layers, such as Jiao YanZu in China, were created to respond to the need for greater school-to-school coordination and interaction. They therefore draw their membership from schools (usually principals), with minimal additional administrative or technical support; these mediating layers differ from others in that their school-to-school orientation dominates their role, rather than the usual school-to-center one. The fourth form of mediating is that of the school systems that have separate sub-structures for primary schools and secondary schools, such as Long Beach's Unified School Districts.

While their form within the system varies somewhat, the functions the mediating layer has played in maintaining system improvement is fairly consistent. They typically have three tasks:

- Providing targeted support to schools.
- Acting as a buffer between the center and the schools while interpreting and communicating the improvement objectives, in order to manage any resistance to change.
- Enhancing the collaborative exchange between schools, by facilitating the sharing of best practices between schools, helping them to support each other, share learning, and standardize practices.

We now look at each of these three roles in turn.

## Providing targeted support to schools

Exactly what support is provided and how each mediating layer achieves this support varies somewhat between systems (Exhibit 30). To understand better how they do this in practice we will focus on the example of Western Cape.

In socio-demographic terms, the Western Cape province is a highly varied province, as are all of South Africa's nine provinces. The Western Cape's 1,500 schools are spread across eight districts, which range from Cape Town's wealthy southern suburbs, to school districts comprising much poorer, historically disadvantaged and densely populated urban communities, to the rural schools of the Cape Winelands District. The Provincial Department of Education decided early on in its journey that it needed to incorporate an approach that was responsive to the wide range of schools' needs across the province. However, it was clear that it could not leave its plans for outcomes improvement as the responsibility of each school – the capacity constraints were too great, student outcomes too low, and the need to improve was too urgent for that. Western Cape was therefore faced with the need to devise an approach that would be responsive to the varied needs of its districts and schools without being laissez faire and completely reactive.

In 2002, the Provincial Department halted a central-run, expert-led process for developing a new curriculum as this was failing to achieve the desired results, and called the district leaders together in order to develop a literacy strategy. They defined three areas of improvement on which each district was required to focus: one, teacher development and support: two, the provision of resources and learning materials: three, research and advocacy. Within this framework, however, districts would be free to adopt different approaches to implementation in response to how their defined their schools' needs. The Western Cape believed that the districts were better positioned than the province to meet its objective in targeting support to the needs of schools and communities. For example, districts could decide how to allocate attention and support across schools, how to configure their team support, and what interaction routines they set with schools. A provincial literacy and numeracy coordinating committee was set up to create accountability and integration for implementation across the province. Rather than diminishing the role of the province with its schools, working through the districts increased the ability of the province's head of the Literacy and Numeracy Strategy to drive forward the implementation. The role of the mediating layer enabled him to stay very hands-on when working with district leaders and in holding districts and schools accountable for their outcomes in third and sixth-grade assessments.

The addition of this mediating layer increased the intensity of interaction between the schools and the center. As the level of support increased significantly, the relationship changed from one ➔

# Exhibit 29:
# The middle layer plays an important role in delivering and sustaining improvement

| | Schools Number | Middle layers Layer 1 | Layer 2 | Layer 3 |
|---|---|---|---|---|
| Aspire | 30 | Areas (3) | | |
| Long Beach | 92 | Level Superintendents (2) | | |
| Boston | 135 | Cluster (9) | | |
| Singapore | 326 | Clusters (28) | | |
| Slovenia | 977 | Municipalities (210) | | |
| Latvia | 982 | Municipalities (118) | Regions (5) | |
| Hong Kong | 1,196 | Regional offices (4) | | |
| Lithuania | 1,415 | Municipalities (60) | Counties (10) | |
| Armenia | 1,452 | Marzer (10) | | |
| Saxony | 1,480 | Regional subsidiaries of education agency (5) | | |
| Western Cape | 1,569 | Circuits (49) | Districts (8) | |
| Jordan | 4,280 | Governorates (12) | | |
| Ontario | 4,423 | School boards (72) | | |
| South Korea | 11,383 | Districts/counties (118) | Provincial offices (16) | |
| Chile | 11,763 | Communes/municipalities (346) | Provinces (138) | Regions (15) |
| Minas Gerais | 17,900 | Municipalities (853) | Regional Departments of Education (46) | |
| England | 24,570 | Local gov. authorities & boroughs (353) | | |
| Poland | 25,769 | Gmiras (Municipalities) (2,478) | Powiats (countries) (379) | Voirodeships (provinces) (16) |
| Ghana | 28,764 | Districts (170) | Regions (10) | |
| Madhya Pradesh | 138,527 | Blocks (~250-350) | Districts (48) | |

Source: TIMSS, PISA, NAEP, national and provincial assessments; team analysis

# Exhibit 30:
# Middle layers provide targeted support to schools

| | Description |
|---|---|
| **Ontario, Canada** | ▪ 2009 changes to Education Act increase School Boards' responsibility for student achievement (in addition to administrative and financial responsibilities)<br>▪ Student Achievement Officers facilitate professional learning communities for Principals within School Boards |
| **South Korea** | ▪ District offices offer training for teachers based on needs of schools within district<br>▪ Implementation of reform often cascaded through provincial offices (e.g. KEDI's ICT reforms in the 1990s) |
| **Western Cape, South Africa** | ▪ Districts each have cross-functional circuit teams that provide improvement support to schools (literacy advisers and curriculum coordinators, as well as administrative support)<br>▪ Districts/circuit identify locally specific issues and develop locally tailored solutions: (e.g. lobbying wine farmers association to allow farm workers leave to visit their children's schools) |

Source: System interviews

of occasional visits from the province or district to one in which a team was housed "on the doorstep of the schools." The tone of the interaction changed too, from one where schools felt "inspected" to one of partnership and support. Every week, circuit teams meet to discuss the school visits, to problem solve the challenges they face, and to draw support from the district as needed, as well as from the province or third party partners such as NGOs and community organizations active in the area.

No one would claim that everything is perfect. The educators we spoke with in the Western Cape were quick to make the point that gains in numeracy have not matched those in literacy, and that third-grade literacy gains have outpaced sixth-grade gains. However, they were also quick to say that this was just a start, and that they were clear in their conviction that they have begun to develop a model that works for their system.

## Acting as a buffer

The mediating layer plays an important role between the schools and the center, amplifying messages that are important to reform while buffering any resistance. The mediating layer can amplify constructive communication by ensuring each school receives and understands guidance from the center, and that the center hears feedback, requests and ideas from schools. It can also buffer resistance to change, resolving the issues that can be tackled locally and highlighting those that the center need to understand and deal with, while filtering out much of the unconstructive noise that always accompanies challenging changes. To see how this works in practice, we look at the role the mediating layer has played in Poland's improvement journey.

As has already been mentioned, in 1999 Poland was faced with the daunting task of opening 4,000 new lower secondary schools. In large part this was to be achieved by closing primary schools and reconstituting them as lower secondary schools. Doing so was contentious. Parents and teachers, concerned about the possibility of their local primary school closing, protested against the change. Each of Poland's 2,500 municipalities was tasked with implementing the restructuring, and was given the flexibility to tailor their approach to each community. Empowering the municipalities to directly resolve communities' concerns allowed the system to engage much more deeply with each community than would have been possible if the Ministry had tried to directly communicate with each community.

Municipalities worked with their communities to build understanding about both the quality and affordability of the potential changes. In consequence, the different municipalities developed different solutions. The municipalities had to tackle practical challenges resulting from the restructuring, such as busing children in order to reduce longer journey times. Certain municipalities made concessions in exchange for the community's acceptance of necessary but difficult school changes. For example, some bartered new community infrastructure, such as the provision of a bridge or road, in exchange for acceptance of the restructuring. By empowering the municipalities in this manner to implement the reforms, Poland dampened resistance while developing solutions that made the changes more palatable to both the schools and their communities.

## Enhancing the collaborative exchange between schools

A third way in which the mediating layer fortifies system improvement efforts is by opening up channels between schools to share learning, standardize practice, and support each other, as is the case for both Singapore's and Boston's school clusters. Singapore established its school clusters in 1997 as forums for principals to share experiences and best practices, and to do some local-level resource allocation. Boston Public Schools created nine geographic school clusters to provide a forum for peer-to-peer support and sharing for principals. Boston's cluster leaders were selected from among highly effective principals so that they would be in a position to mentor the other principals in the cluster. With the formal connection at the level of the school principals, the school clusters also became a network for inter-school interaction between both teachers and students.
The mediating layer played a similar role in this respect in both China and Western Cape. In China, Jiao YanZu standardized practices across its schools within each district and provided a forum where subject leaders could share with their peers within

the district. In Western Cape, its district level Literacy Coordinators were close enough to the schools to identify what was working and what was not; they could disseminate what they found to the senior provincial leadership where necessary. The coordinators met routinely with their circuit and districts teams, and quarterly with the Provincial LitNum Coordinating Committee. In these meetings they developed solutions to the challenges the schools were facing, and shared approaches about what they found worked effectively in schools. One Western Cape Literacy Coordinator describes how this worked in practice: "If I hear about something good going on in a school, I will go visit it. After I see it, I may tell the Provincial Deputy Director General about it and I will share it with the Provincial Coordinating Committee. That way it gets to the top and to others who can use it."

Across all the systems we studied, despite their differences in structure, the mediating layers were effective in opening up channels for communication, sharing, support, and standardization between the schools themselves and from the schools to the center.

## Architecting Tomorrow: The CPU

No computer system works without regulation and continuity. This continuity is dependent in part on the operating system and how it is interpreted through the user interface – but dig deeper and it becomes apparent that the efficacy of the computer system is dependent on the proper functioning of the central processor. It is the central processor that ensures that each function happens in the same manner this time and the next time as it did the last. In school systems it is the system leaders who ensure such continuity: they do so by ensuring that both the explicit and implicit pedagogical aspects inherent in the school system are transferred to future generations of leadership.
The most successful examples of leadership continuity we encountered come from systems that are able to develop their future leaders from within their system. For 30 years Singapore has had a story of sustained improvement, changing tack to set course towards new horizons as times change, but never stopping, never doubling back to unwind the past, always moving forward. It also possesses the most structured approach to identifying and developing future system and school leaders that we encountered among all 20 systems.

Using its system of career paths we described earlier in this chapter, Singapore systematically identifies and develops talented educators for leadership positions from within the school system. All educational leadership positions up to the level of Director-General Education are considered professional positions and are part of the teaching career structure. All promising teachers are put onto this career track, thereby develop a pipeline of school leaders. Teachers with the potential to become principals are identified at an early stage and appointed to middle-leadership positions in schools, as subject or level heads or as heads of department. To better prepare them for their management responsibilities, they attend a full-time four-month milestone program (Management and Leadership in Schools) at Singapore's National Institute for Education. Educators considered ready for the next level of leadership appointment will be interviewed for appointment as vice principals. Vice principals attend a six-month Leaders in Education program, which has an executive orientation similar in scope and intensity to executive courses in business schools, but with a focus on education.

Since the 1980s, Singapore has paid particular attention to the professional development of principals and continues to evolve the support and apprenticeship they receive. Newly appointed principals are paired with more experienced ones under a mentoring scheme started in 2007. They also receive "CEO-style" development programs. Experienced principals are given sabbatical opportunities, and top principals can become Cluster Superintendents, as a first step to system leadership.

Over the past twenty years Long Beach Unified School District has become a model for urban public school system transformation in the United States and provides a further model of how to build continuity in leadership. During these two decades, Long Beach has had just two Superintendents: ➔ Carl Cohn, and his successor Chris Steinhauser,

# Exhibit 31:
# System stability can come from either the political or the strategic leader

**Strategic leaders**

| | Strategic leader | Tenure Years | Political leaders |
|---|---|---|---|
| Armenia | Karine Harutyunyan | 15[1] | ▪ Since 1995, Armenia has had ten Ministers of Education |
| Western Cape | Brian Schreuder | 10[1] | ▪ Since 2000, there have been five MECs (Member of Executive Council) for Education for the Western Cape |

**Political leaders**

| | Political leader | Tenure Years | Strategic leaders |
|---|---|---|---|
| Singapore | Lee Kuan Yew | 31[2] | ▪ Multiple Permanent Secretaries and Ministers of Education throughout tenure, including between 1980-90 |
| Ontario | Dalton McGuinty | 7[1] | ▪ Since 2003, three Deputy Ministers and two Chief Student Achievement Officers |

1 Was still in position at time of research
2 Lee Kuan Yew was Prime Minister from 1959 to 1990, although the primary focus of this report is from 1980 onwards.

Source: System interviews

who was Carl's deputy before that. As a result, there has been remarkable consistency in the architecting of practices and mindsets. Signature leadership practices such as "Cookies with Carl" later become "Coffee with Chris," both of which were informal meetings with school staff and the community to discuss concerns and needs. Though there have been changes in the system's priorities and approach, these have been the result of system evolution, not revolution. Long Beach's culture of consultation, the collective ownership of its schools, data-driven decision making, and the focus on what students learn rather than what teachers teach are all deeply embedded in the system.

Not every system is blessed with the advantage of Long Beach's leaders' long tenures. Lithuania demonstrates how a system can create continuity even when tenures are short. Following the dissolution of the Soviet Union, its first Minister of Education and Culture was Darius Kuolys, who held this position from March 1990 until December 1992, during which time he designed and launched Lithuania's school system strategy. The subsequent minister asked Kuolys to stay on as an advisor, which he did until April 1993. During this period (1991-93), Kornelijus Platelis served as Deputy Minister of Education and Culture. Later, Platelis served as Minister of Education from May 1998 through November 2000, while Kuolys served in parallel as senior advisor to the President on social issues, including education. Throughout this period, other members of the ministry rose to more senior positions. To take just one example, the current Director of Education (in charge of the school system) was then a member of the Kuolys strategy design team and has held multiple roles overseeing the strategy implementation over the years. In this manner, Lithuania's system leaders have been able to groom the future leadership while still in office, providing continuity for the system.

Though Ontario's system improvement journey has a shorter history, it has also followed a similar pattern to Lithuania in identifying its future system leaders from within its ranks. Kathleen Wynne, Ontario's Minister of Education from 2006-10, had previously been parliamentary assistant to the former Minister of Education, Gerard Kennedy, from 2004-06. Mary-Jean Gallagher, Assistant Deputy Minister of Student Achievement, was formerly the director of a district school board before taking over the leadership of Ontario's literacy and numeracy strategy from Avis Glaze. One priority during Mary-Jean Gallagher's tenure has been to incorporate the Literacy and Numeracy Secretariat into the Ministry of Education, transforming it from a reform strategy into a permanent and core function of the ministry. Other leaders, such as the Kathleen Wynne's successor as Minister of Education, Leona Dombrowsky, came from outside Ontario's school system, but had previously been part of Premier Dalton McGuinty's government and so were already familiar with the system's priorities and its approach to managing improvement. Michael Fullan, a prominent educationalist and academic based at the Ontario Institute for Education Studies, served as a close advisor to Premier Dalton McGuinty throughout the journey, also providing a layer of continuity.

Other systems have relied upon a single anchor in the form of a political or strategic leader (Exhibit 31).[35] For example, Armenia's strategic leader's tenure has been fifteen years so far, while ten ministers of education have come and gone during this time. In England, during Tony Blair's tenure as prime minister, there were five education secretaries and two deputy ministers. So despite the frequent turnover of other system leaders, these systems have been able to provide leadership continuity.

Each of these improving systems has written a consistent story of improvement by ensuring the leaders who shepherd the system share the experience and ownership of the system's pedagogy. When pivotal leadership roles are filled, these systems have usually been able to identify leaders from within their system with the required capabilities and experience to fill them. That they were able to find such leaders is no mere accident – these systems deliberately set about architecting for tomorrow's leaders. As a result, the improvement journey of these systems has been evolutionary – not halting, nor inconsistent, and not repeatedly disrupted. ■

# Ignition

School systems that have successfully ignited reforms and sustained their momentum have all relied on at least one of three events to get them started: they have either taken advantage of a political or economic crisis, or commissioned a high-profile report critical of the system's performance, or have appointed a new, energetic and visionary political or strategic leader.

The role of new leadership is a common and particularly important pattern in igniting school system reforms, occurring in all of the improving systems we have studied. The evidence suggests that leaders that are successful in directing a system's improvement journey are characterized by taking advantage of the opportunity afforded by their being new to the role, in following a common "playbook" of practices, and in their longevity, having a much lengthier tenure than is the norm.

## Getting going

Having got this far in the report, many a system leader might well be asking, "Well, where do I start" or, "What do I need to get the process started?" This chapter sets out to help answer these questions.

School system leaders often feel that the task of system improvement is too big for their limited political mandate, resources, and authority. Even school systems that manage to launch a reform effort often find that the momentum quickly peters out. So what can a system leader do to light the fire and keep it burning in the years to come?

Our research of the 20 improving systems found that there are three circumstances that regularly "ignite" school system reform (Exhibit 32):

- Political and economic crises: crises of grand proportions have often been credited with jolting a change in behavior across multiple domains, ranging from national political leadership to business and cultural practices. We find that education is no different, with regime changes, risks to nationhood, and economic crises sparking school system reform efforts designed either to mitigate the potential ill effects of the crisis or to take advantage of new opportunities. In the word of one Asian school system leader, "Anything that affects the size of our rice bowl leads automatically to reassessing the skills of our nation's youth."

- A high profile and critical report about system performance: public reports regarding poor student outcomes are another powerful source of impetus for igniting school system reform. As one East European system leader noted bluntly, "Sometimes you need to drop a bomb in the system to get people to move." In some cases, international assessment results have played this role; in others, system leaders have commissioned a third-party review of the system, knowing full well that it would result in a strongly negative message about the current system performance. These leaders then broadcast this report widely, both to shock the public and to build as much external pressure as possible to force the hand of the government into launching system reform.

- New political or strategic leaders: as with any institution, a new leader fresh to the system has plausible deniability for past performance and so is in a good position to be able to reset relations with critical stakeholders. We find new leaders present in all of our sample school systems; these new leaders have come both from the political field (e.g. Prime Minister, President) and the strategic one (e.g. Minister of Education, Superintendent of School Districts). Clearly, being "new" is only an advantage for a leader when affiliated with other personal characteristics, such as energy and vision, as the leaders still have to take advantage of the opportunity afforded by their interjection into the system. Our research shows that upon their appointment, successful new leaders, irrespective of the system's context, tend to follow a similar "playbook" of actions designed to ensure that the sparks of reform catch alight.

In several systems, more than one of these circumstances was present to light the fires of the improvement journey; this was the case for Singapore in the late 1970s, Long Beach in the early 1990s, and Poland in the late 1990s. Of the 20 systems that are the subject of this study, 15 saw the conjunction of two such circumstances present at the ignition of their reforms.

Of the three sources of ignition described here, the most prevalent – and therefore the most important – is the impact that a new political or strategic leader can have on a system: in all 20 of the 20 systems we studied a new leader sparked the fires of reform. Moreover, we also see a pattern that when new leaders are drawn from outside the system they are much more likely to break with the past. Once installed, successful leaders of improvement journeys have another thing in common – staying power. We find that the median tenure of the leaders in our studied systems is six years for strategic leaders and seven years for political ➔

# Exhibit 32: Each improving school system employed at least one type of ignition

| System | Political/Economic Crisis | Publication of Critical Report | New Leader |
|---|---|---|---|
| Singapore ('79) | × | 1978 Goh Report | ✓ |
| Singapore ('97) | ✓ – 1997 financial crisis/regime change | × | ✓ |
| Hong Kong ('80) | × | × | ✓ |
| Hong Kong ('00) | × | × | ✓ |
| South Korea ('98) | 1997 financial crisis | × | ✓ |
| Boston ('95) | ✓ – Referendum: Mayor to appoint school board | × | ✓ |
| Ontario ('03) | × | × | ✓ |
| England ('97) | × | × | ✓ |
| Saxony ('90) | ✓ – Fall of Berlin wall; reunification of Germany[1] | × | ✓ |
| Slovenia ('92) | ✓ – post Soviet | × | ✓ |
| Aspire ('98) | × | × | ✓ |
| Poland ('97) | ✓ – post-Soviet | × | ✓ |
| Latvia ('90) | ✓ – post-Soviet | × | ✓ |
| Lithuania ('90) | ✓ – post-Soviet | × | ✓ |
| Long Beach ('92) | ✓ – 1992 loss of 2 biggest employers | × | ✓ |
| Armenia ('95) | ✓ – post-Soviet | × | ✓ |
| Chile ('91) | ✓ – post military rule of Pinochet | × | ✓ |
| Minas Gerais ('03) | × | × | ✓ |
| Madhya Pradesh ('05) | × | 2005 ASER report | ✓ |
| Western Cape ('01) | × | 2001 GTZ report | ✓ |
| Ghana ('03) | ✓ – 2001 shift to democratic government | × | ✓ |
| Jordan ('00) | × | × | ✓ |

1 While PISA was considered a shock for Germany, interviews indicated that it was not the motivation for the Saxony reforms

Source: McKinsey & Company interventions database and system interviews

leaders. Stable leadership enables continuity in the reform agenda. This is in stark contrast to the norm: for example, the average tenure of system leaders is nearly three years for Superintendents of urban school districts in the U.S., and two years for Education Secretaries in England, as well as for education ministers in France.

School system leaders around the world pursue reform with great dedication, energy, and passion. However, meaningful school system reform can be hard to achieve – even putting aside political and structural constraints. The danger is that the flames never really catch hold, or die away too soon because leaders find themselves mired down in the day-to-day detail. Later in this chapter we describe the patterns identified in our research of how leaders use a playbook to ensure that the path of change stays true. First, we look at how each of the three circumstances listed here contributed to lighting the fires of school system reform.

## Never waste a good crisis

The dissolution of the Soviet Union in the early 1990s marked a sharp disjuncture for all the nations under its sphere of influence. Armenia, Latvia, Lithuania, Poland, and Slovenia, all improving systems in our sample, chose to launch school reforms in the period that followed. Some of these reforms were the natural consequence of this disjuncture, such as replacing Soviet history textbooks with new ones that focus on their own national history, or in introducing fresh topics such as civics and democracy. These systems, however, went further, taking advantage of the new openness to launch a bold sweep of reforms that included decentralizing school system management, revising the school model, optimizing the number of schools and staff, and changing the school funding model. One Lithuania system leader notes: "A period of euphoria had taken hold in the country, with teachers and principals feeling like they were personally contributing to building the nation and to driving system change at high speed. Even changes that hurt certain parties were considered acceptable for the sake of the greater good of our children and nation."

In Poland, Prime Minister Jerzy Buzek announced broad reforms in 1998 that cut across four major areas: education, healthcare, government administration, and pensions. As has been described in Chapter 1, this announcement resulted in the 1999 school reforms, led by the Minister of Education, Miroslaw Handke. The signature initiatives of these reforms were the extension of general education by one year (thereby requiring the reconstitution of 4,000 primary schools in one year and reopening them as lower secondary schools), and the decentralization of the administrative and financial decision rights over schools to municipalities. They built on strong public sentiment in favor of decentralizing as far as possible to the local level, to "give power back to the people," in reaction to the centralizing tendency of both martial law and communism. This drive for decentralization manifested itself in multiple ways, including in enabling principals to choose which teachers they wished to hire, in allowing teachers to choose which curriculum they wished to use from among 100+ private providers, and in empowering communities to create their own schools.

Armenia found itself more challenged than most when it came to national survival. Following independence from the Soviet Union in 1991, the country was plunged into war with Azerbaijan. The cost of the war, in combination with Armenia's landlocked status, caused it to fall into deep financial crisis. Its funding level dropped from USD 500 per student per year during the Soviet period to just USD 24 per student a year in 1994. As a result, schools had to be shut down during the winter for several months due to the lack of power for heating. Teacher salaries fell to only USD 5 a month in 1995, the equivalent of unemployment benefits. Finding itself in such dire straits created pressure on the system to significantly restructure its funding model and to optimize the number of schools and teachers in the system; in parallel, it also upgraded its curriculum in core subjects in order to raise student skill levels.

Likewise, Long Beach also experienced a hard-hitting economic crisis that precipitated its schools reform. In the early 1990s the district witnessed an economic depression, characterized by its system leaders as being worse for the state than ➔

today's international financial crisis. The district had lost its two main employers and the 35,000 jobs that went with them. Furthermore, with escalating gang violence, Long Beach was experiencing a mass exodus of wealthier families into neighboring suburbs. Superintendent Cohn took office in the wake of the 1992 riots, and worked with the community to start the Seamless Education System – a collaborative venture between higher education institutes and the school district designed to improve school system performance.

Crises have played a similarly strong role in catalyzing Asian systems' education reforms. Singapore's abrupt independence from Malaysia in 1965, in the midst of economic and social turmoil, presented Lee Kuan Yew with the opportunity to declare that skilled human capital was Singapore's only path to prosperity, thereby sparking decades of deep-rooted school system reform. In 1997, Hong Kong's transfer of sovereignty led it to take multiple steps to increase its self-reliance, including shoring up its school system to support its economic competitiveness within the region and more broadly.

Across our 20 sample systems, we find that regime changes, risks to nationhood, and economic crises have all sparked systems to embark on a journey of improvement. System leaders have been adept at successfully tapping into public sentiment, explicitly making the link between the crisis at hand and the requisite school system reforms. In the words of one Asian system leader, "It is not just about riding a crisis and trying to plug school system reform into it. It is about making clear that school system reform is essential to surviving the crisis itself."

## Nowhere to hide

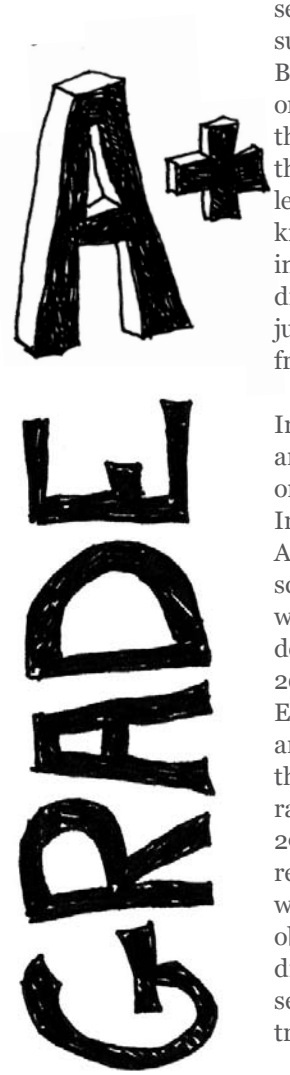
The publication of a high-profile public report, containing overwhelming evidence of low student achievement, has also on occasion been the catalyst for reform. The evidence contained in high-profile reports can ignite debate and lead the public to hold the government accountable for its response to the failings. International assessments, such as PISA and TIMSS, have also caused vigorous national debate about student skills, stimulating changes in the direction the school system is taking. For example, publication of the PISA results in late 2001 led to the so-called "PISA-Schock" in Germany. While the German public's perception of its education system was still high, the results ranked it among the bottom third of countries participating in PISA. Germany's PISA-Schock led to widespread media coverage and intense national debate, giving significant momentum to school reforms across Germany.

Savvy leaders often take advantage of critical external and internal reports on student outcomes to help them push through their reform agenda. For example, in May 2005, an Indian NGO named Pratham published its first "Annual Survey of Education Report" (ASER), offering comprehensive data across India on student performance in literacy and numeracy. The ASER was troubling for India overall, and particularly for the state of Madhya Pradesh – only 57 percent of Madhya Pradesh's first- and second-standard students (ages six to seven) could read at their grade level (compared to the national average of 70 percent) and only 58 percent of its students in standards three through five (ages eight to ten) could read a first-standard text or above (as compared to national average of 67 percent). The report's finding received widespread media coverage, triggering the political leadership in Madhya Pradesh to focus on school system reform. The newly elected state government, led by a new Chief Minister, Shivraj Singh Chouhan, launched the "Learn to Read" program in 2005 in order to improve state literacy and numeracy outcomes for every student from the second to fifth standards. The program addressed the root cause of underperformance uncovered by the ASER report. For the first time in the state, the reforms concentrated on student outcomes and skills instead of the state's traditional focus on inputs (e.g. the numbers of teachers recruited and schools built). Its ambition was that all students should reach the prescribed reading and arithmetic levels.

In some cases, a report that has been commissioned internally by an education ministry has had similar impact. As already describe in Chapter 2, in Singapore, the 1978 Goh Report, named after the Minister of Education, Dr. Goh Keng Swee, was seminal in the country's education history, leading to a major overhaul of its primary and secondary schooling system. The Goh report identified three

main challenges for the country: the low rate of progression of students from primary to secondary level, low literacy achievements, and the poor and uneven quality of instructional materials. It also noted the low morale of teachers and argued that addressing this situation would be critical in attracting and retaining the talent necessary to address these challenges.

The key recommendations of the Goh report shaped Singapore's education system for the next thirty years. The most prominent of these recommendations was that its students should be "streamed" into ability groups. Implemented in January 1979, the rationale for streaming was based on the report's findings that Singapore's then universal "6-4-2 curriculum" (six years primary, four years secondary, and two years pre-university) was not sufficiently customized to differing student abilities. By streaming its students into three groups, based on their aptitude, and providing each group with the appropriate curriculum, the report contended that students would be better placed to be able to learn at an appropriate pace, to acquire appropriate knowledge and skills, and would be encourage to stay in school longer. Indeed, at the primary level, the dropout rate fell from six percent in the late 1970s to just 0.5 percent in 1997; at the secondary level it fell from thirteen percent to 3.3 percent.




In yet other cases, the school system commissions an external review of its own performance in order to build justification for a reform effort. In 2001, the province of Western Cape in South Africa worked with GTZ to conduct a review of its school system. Western Cape's 1,500 schools face wide socioeconomic diversity and the legacy of deeply unequal education. As part of this review in 2002, GTZ and the Western Cape Department of Education conducted literacy tests for its third-grade and sixth-grade students. The results shocked both the educators and the public, showing that the pass rate was only 36 percent in the third grade and just 29 percent in the sixth grade; not only were these results surprisingly poor, but they got progressively worse as students got older. One system leader observed, "Some people protested it, but we in our district were thankful. It proved what we were seeing and got everyone's attention." In a system traditionally focused on ➔

## Exhibit 33:
## New strategic leaders were present in all reforms we studied, while new political leaders were present in half of them

| System[1,2] | New strategic leader? | Name(s) – Strategic leader | New political leader? | Name(s) – Political leader[3] |
|---|---|---|---|---|
| 1a Singapore ('79) | ✓ | Goh Keng Swee[4] | | |
| 1c Singapore ('97) | ✓ | Lim Siong Guan/Chiang Chie Foo | | |
| 2a Hong Kong ('80) | ✓ | YT Li | | |
| 2b Hong Kong ('00) | ✓ | Fanny Law | | |
| 3 South Korea ('98) | ✓ | Lee, Seok-Hee | ✓ | Kim, Dae-Jung |
| 4 Boston ('95) | ✓ | Tom Payzant | ✓ | Thomas Menino |
| 5 Ontario ('03) | ✓ | Ben Levin | ✓ | Dalton McGuinty |
| 6 Saxony ('90) | ✓ | Wolfgang Nowak | | |
| 7 England ('97) | ✓ | Michael Barber | ✓ | Tony Blair/David Blunkett |
| 8 Slovenia ('92) | ✓ | Slavko Gaber | ✓ | Milan Kucan |
| 9 Poland ('97) | ✓ | Miroslaw Handke | ✓ | Jerzy Buzec |
| 10 Latvia ('90) | ✓ | Andris Piebalgs | | |
| 11 Lithuania ('90) | ✓ | Darius Kuolys | | |
| 12 Long Beach ('92) | ✓ | Carl Cohn | | |
| 13 Armenia ('95) | ✓ | José P Arellano Marín | ✓ | Eduardo Frei Ruiz-Tagle |
| 14 Chile ('94) | ✓ | Karine Harutyunyan | ✓ | Aécio Neves da Cunha |
| 15 Minas Gerais ('03) | ✓ | Vanessa Guimaraes | | |
| 16 Western Cape ('01) | ✓ | Brian Schreuder | | |
| 17 Ghana ('03) | ✓ | Ato Essuman | | |
| 18 Jordan ('00) | ✓ | Khaled Toukan | ✓ | King Abdullah bin Al-Hussein |

1 Aspire was not included in this analysis as by definition charter schools are new entities with new leaders
2 Madhya Pradesh was not included as interviewees did not have consensus on the strategic leadership role.
3 These are political leaders indicated by interviewees to be highly engaged in steering educational reform. It excludes political leaders who were described to have more supported reforms without strong and active steering, even if they were new leaders at the beginning of reform
4 Although a Minister, Goh Keng Swee was identified by interviewees to have provided the strategic leadership for Singapore's 1980 reforms

Source: McKinsey & Company interventions database, system interviews

university matriculation exams, the evidence generated by the reports findings focused attention sharply on primary school learning for the first time, and particularly on improving literacy. This motivated Western Cape to launch its new literacy strategy in 2003 (described in Chapter 1 of this report), which was designed to improve the quality of classroom teaching, strengthen school management, and provide adequate resources to students.

The publication of a report exposing the frailty of a school system holds inherent risk, of course: political leaders often view with trepidation the negative messages they contain and can all too easily shy away from placing them in the public domain. However, examples among the 20 improving school systems studied here demonstrate that searing evidence of poor achievement can be harnessed effectively to spur reform. It can serve to put pressure on the parts of the system that are comfortable with the status quo. As one Eastern European leader observed, "Being deliberately and uncomfortably conspicuous can sometimes be exactly what a system needs to alter its direction."

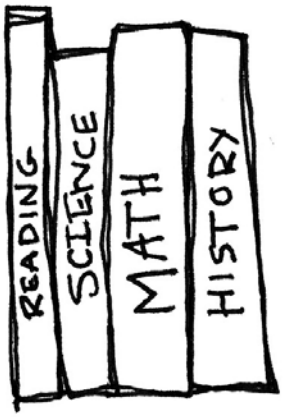
## Entering stage right

Of the three circumstances that ignite reform and start a system on its improvement journey, the introduction of a new political or strategic leader is by far the most commonly observed across our sample systems. Indeed, it is universal: all the systems in our sample started their improvement journeys under new strategic leaders, and half had new political leaders (Exhibit 33).

Who are these new leaders, and where do they come from? They can be political leaders (e.g. President, Prime Minister, Governor) or strategic (Minister of Education, Superintendent of Schools), or both. What typifies either type is their focus. Ontario's Dalton McGuinty, for instance, is widely known as the "Education Premier" because of his persistent focus on education as his number one priority. Similarly, the German state of Saxony has been led by center-right coalitions over the past nineteen years, and during this time all three of its minister-presidents have regarded education as their top priority.

The political leader is understandably motivated to bring in a new strategic leader to jump-start a reform process. When a system's leaders have lost personal credibility, or have strained relations with stakeholders, or have overseen a period of declining or stagnating student performance, the presence of a new school system leader can help change direction. Not only do incoming strategic leaders have the ability to break with the status quo, they potentially come with a "clean slate," which they can use to improve relations with stakeholders.

For the political leader, the question is whether to appoint a new strategic leader from within the system or bring one in from the outside. The relative efficacy of choosing an insider or an outsider depends on the degree of disruptive change required in the system. For example, when Prime Minister Lee Kuan Yew installed Dr Goh Keng Swee as Singapore's new Minister of Education in 1978, the latter's mission was to address very challenging problems: almost 30 percent of the country's primary students were not progressing to secondary school, and the government wished to raise the efficiency of the Ministry of Education in managing the school system. Goh was a complete outsider to the Ministry of Education, previously serving as Minister of Finance and Minister of Defence. In contrast, Long Beach has had a long tradition of hiring from within the system to fill its superintendent role. Even Carl Cohn, who became Superintendent during a challenging period for Long Beach, was considered to be an "insider with an outsider's perspective" because he had initially worked in Long Beach, then left for a period, before returning to the system as superintendent. Chris Steinhauser, who had been Cohn's Deputy Superintendent, succeeded his superior in 2002. With Steinhauser's appointment, Long Beach had returned to its previous tradition of hiring from within; it did so because the district was by then on a positive improvement trajectory and many felt its most pressing requirement was leadership continuity. ➔




## The new leader's playbook

The appointment of a new leader provides a disjuncture for the system, but it is what these leaders do with this opportunity that determines their ultimate success or failure. Our research shows that once they enter their positions, successful new leaders follow a common "playbook," irrespective of system performance level, culture, or geography. This playbook comprises five steps.

### Step 1: Decide on what is "non-negotiable"

The first major challenge for a new leader is to decide which levers they should use to improve the system. For each of these levers, the leader must then define a small set of fundamental rules or actions: we call these "non-negotiables," borrowing a term that is often used by Long Beach system leaders. These non-negotiables become the anchor points of the system reform. Successful leaders are vigilant in ensuring that there is little or no compromise in their execution, whereas compromise on other reform aspects is often plentiful. As one Lithuanian system leader observed, "A reform is like a big bowl of soup. The cook is stirring the soup, but if the spoon is too short, you only mix the surface. The spoon must be long enough to reach the meat and potatoes that lie at the bottom." The "meat and potatoes" are the non-negotiables.

Two of Long Beach's non-negotiables were student achievement standards and professional development. As one leader stressed, "Our main objective was to break down the silos and be very clear that the central offices exists to support the schools ... grabbing hold of standards and professional development was the only way to make that happen."

Likewise, when Slovenia's new Minister of Education, Slavko Gaber, took office in 1992, he focused on resources, curriculum, and professional development as the non-negotiables. He was also committed to the way in which these changes would come about, "The most important part was that teachers were involved... that teachers had the feeling they were asked." For example, he allowed the consultation with teachers on curriculum standards and design to span four years, refusing to succumb to political pressure to end the debate even when the deadline for announcing the new curriculum loomed close. Instead, he went to Parliament and asked for more time to continue the teacher consultation process. Three months later, Gaber had successfully resolved with teachers the issues pertaining to time allocation by subject, and announced the new curriculum reform in 1996. Similarly, in Poland, the introduction of 4,000 lower secondary schools in one year was their non-negotiable, while in Minas Gerais it was that each school should achieve its literacy improvement target. Each system must select the lines that cannot be crossed.

### Step 2: Install capable and like-minded people in the most critical positions

When Goh entered Singapore's school system in 1978, he brought a team of seven system engineers with him to replace some of the ministry's

educators with people from outside the system. This move was the first of its kind in Singapore's Ministry of Education. The seven filled some of the most sensitive roles in the Ministry, including the Permanent Secretary, Director of Education, head of the Curriculum Development Institute of Singapore, and head of Information Systems. Says one senior staffer during this period, "Goh had full support of the Prime Minister to shake things up ... even though this was a big slap to the ministry professionals at the time." In effect, Goh was seeking to control the key inputs of students, data, and teachers to implement the government's streaming strategy and to raise ministry performance. Subsequent Ministers of Education in Singapore did not make similar sweeping replacements, though most have rotated in a small number of senior civil servants who have a strong implementation track record. For example, the planning function is often led by one of these "outsiders" to ensure appropriate strategy and management experience, as well as to balance weight of educator. Says one senior ministry leader, "Planning is the brain trust that pulls things together ... it is only as good as the people in it and it has to be focused on the priorities of the Permanent Secretary."

This is not to give the impression that system reforms are always led by outsiders: apart from the example of Long Beach, already cited, when Ontario and England made literacy and numeracy drives their flagship reforms, the individuals selected to lead both these respective initiatives were viewed as highly skilled and trusted insiders.

## Step 3: Engage with stakeholders

A Hong Kong school system leader observes, "A school system is like a typhoon: students are in the eye of the typhoon and the adults are running around ... it is very easy to descend into chaos." System leaders have found that engaging with parents, teachers, principals, and community leaders is essential for managing the direction and pace of this typhoon. Underpinning their communication efforts is the belief that stakeholder engagement serves to depoliticize education. Consultation not only airs grievances and concerns about the system's potential direction before policy decisions are taken, but it can also help balance the power and influence of the various stakeholder groups by giving each a platform to air their views. As a consequence of such involvement, some systems have been able to register dramatic gains: for example, Ontario went from a position where 24 million student days had been lost due to labor union disputes during the tenure of the previous premier, to one in which, during McGuinty's first two terms in office, not a single day was lost to disputes.

Several of the system leaders we interviewed began their tenure with a tour throughout their territory in order to listen to their stakeholders and hear their concerns directly. In some cases this has eased past relationship tensions and helped to make a fresh start. Such a tour is only the first step in improving relations, of course; to be successful, it has to be followed by continuous communication. There are differing approaches to this. For example, Singapore's First Permanent Secretary of the Ministry of Education in 1997, Lim Siong Guan, required all the top team members of the Ministry of Education to visit fifteen schools per year, "To hear the voice of the teachers and principals." In Long Beach, Superintendents Cohn and Steinhauser constantly consulted the community. From "Cookies with Carl," when Carl Cohn was Superintendent, to "Coffee with Chris," when Carl Steinhauser became Superintendent after him, Long Beach has held monthly forums whereby parents, teachers, and community leaders can speak to the Superintendent about their concerns and discuss the steps the district is taking to improve the school system. Similarly, former Boston Superintendent Tom Payzant regularly met with parents and communities in their local churches in order engage with their concerns. There is a broader and very pragmatic reason why systems engage with parents in particular – to enlist their support in improving their children's outcomes. Students spend less than fifteen percent of their childhood in school, while more than half their time is spent at home and in the community (Exhibit 34). Above and beyond the parent's support for their child's learning at school, their stewardship of this time is therefore critical and high impact. Some systems are able to rely on deeply embedded social values to achieve this goal: for example, in Korea during the 1960s, the cultural norm was that families would "sell the farm" in order to provide for their children's education. ➔

Even today, families in Korea still figuratively sell their farm for their children's education, spending a significant share of their income on private tutorials.

In many other systems, however, parent support has been cultivated. For example, as has already been reported, Western Cape takes numerous measures to engage parents from low-income rural districts. It conducts a road show to share system performance data with parents and to let them know about subsequent district activities; its literacy coordinators bring together illiterate parents to jointly "write a story" that the parents then memorize and "read" to their children; the district office even goes as far as to contact farm owners to encourage them to allow farm workers time off to meet their children's teachers.

Ontario has established a Parents Reaching Out (PRO) fund, whereby school councils and parent organizations can apply for grants to undertake activities designed to support student learning at the school, regional, or province level. One Ontario parent association member notes, "It is less useful to ask parents to get involved in helping the system through parent councils and volunteering at school. It is much more important for us to support parents in their role as parents to support their child's learning." Parent projects funded by PRP can be up to 1,000 dollars at the school level, while those at regional or province level can be up to 30,000 dollars. Since 2006, Ontario has funded more than 5,500 school council PRO grants and more than 200 regional projects, with a total investment of more than 10 million dollars.

Boston has created a Parent Outreach Officer position in its schools to support parents, while in Aspire's charter schools, parents must commit to thirty hours each year to participate in parent-teacher conferences and Saturday School, which is oriented to the whole family. At the end of the year, principals recognize those parents that have given 30-50 hours of their time to the school during the year.

The experiences of these systems suggest that school systems can create a positive experience in engaging with parents, in relation to their children's schooling; the more parents engage with the school, the more likely it is that they will actively support their children's education and help raise performance accordingly.

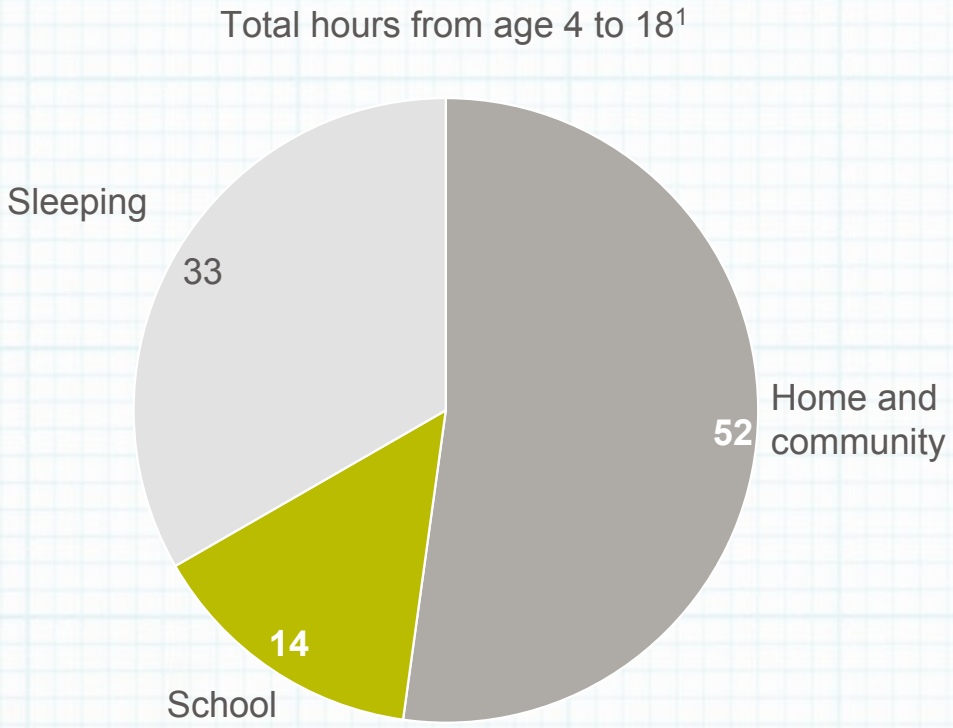
### Step 4: Secure the resources for non-negotiables

"Money is the most important tool of influence," says one Asian system school leader. For system leaders to achieve the non-negotiable system improvements in cases when budgets are tight or insufficient, they are often required to either redirect existing human and financial resources to these activities, or to inject new resources drawn down from the government or private donors.

England took the reallocation path. In accordance with its priority on raising literacy and numeracy, it reallocated its human and financial resources from the existing schools budget to create a supply of literacy and numeracy coaches, teacher training programs, and regional directors. Approximately eighty million pounds per year were allocated to the program between 1998 and 2001. This was at a time when total expenditure on primary teachers' pay was about seven billion pounds, a ratio of 1:87. The program was founded on the notion that relatively small sums of money spent on teacher development could have a major impact on raising outcomes when connected to a clear strategy. The "numeracy program" (1997-2001) appointed fifteen regional advisors for primary schools, and money was given to the Local Authorities to appoint "maths consultants"; the consultants were in essence high-performing math primary school teachers. The head of England's numeracy program trained her regional maths advisors, who then trained the maths consultants appointed by the Local Authorities, who in turn trained the primary school teachers in their local schools. In addition, the center created a five-day course that was attended by thousands of primary teachers; this was both a math refresher course and a showcase of techniques for teaching primary mathematics more effectively (e.g. fractions, long division).

Minas Gerais, Brazil's third largest state, also took the reallocation path in implementing its priorities. The reform's objective was to move the system from an outcome in which, in 2006, only 49 percent ➔

## Exhibit 34:
## Students spend less than 15% of their childhood at school




1 Based on 365 days each year; 8 hours sleep each night; 14 years of schooling (180 seven-hour days each year)

Source: McKinsey & Company

of eight-year-olds were reading at the basic proficiency level, to one in which 90 percent would do so by 2010. This objective translated into having to ensure that 11,400 students would meet the proscribed proficiency level during a four-year period. A critical challenge for the program, therefore, was how to reach these students across the state's 2,450 schools in 853 cities and towns. To achieve this, the state did two things. Firstly, it recruited 46 new people for a central team, all of whom had experience in literacy programs: one for each of the state's regional departments. Secondly, it also employed 300 new people in its regional departments to support the effort. The central team served as trainers for the regional departments and their staff, who then supported the schools and teachers in their respective localities. All the new team members were funded by funds reallocated from existing Department of Education budgets.

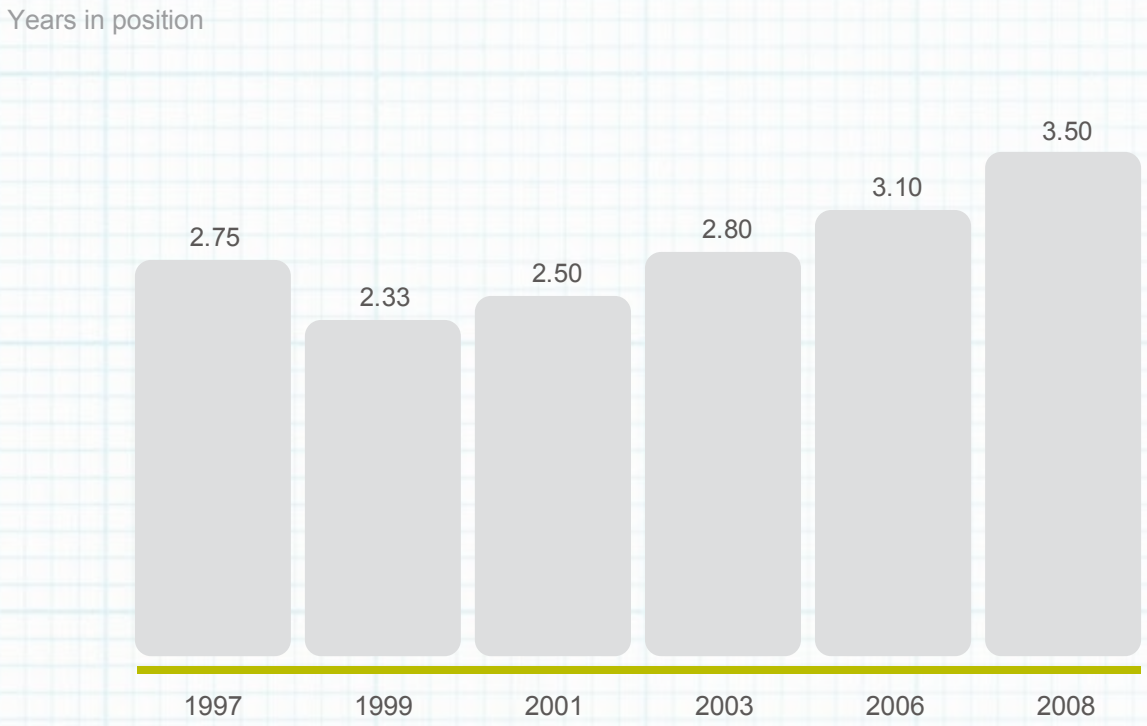
Hong Kong, in contrast, injected new funds into its school system, but relied on the private sector to do this. As Chinese refugees flooded into Hong Kong post-World War II, raising the local population from half a million to three to four million in just a few years, the government was under pressure to provide schooling for an ever-increasing number of students. To meet this demand, Hong Kong established so-called "matchbox schools" in each new public housing building. At first these were set on the rooftops of public housing, then on the ground floor, then they took up the whole top floor, next they were constructed as annexes, and finally as separate buildings. Schools were built at such a rapid rate that literally one rooftop school went up every two weeks. Such a high pace of building construction resulted in the Hong Kong government facing serious human and financial resource constraints. In consequence, they decided to enable the private operation of schools, whereby a private entity could apply to open a school and, if approved, the government would provide eighty percent of capital expenditure and one hundere percent of operational expenditure. By the late 1960s, the overwhelming majority of schools were operated by sponsoring bodies, such as charities and Church missionary societies. To this day, Hong Kong's school system is characterized by being privately operated and publically funded.

### Step 5: Get "early wins" on the board quickly

"If you say you're going to do something, you had better do it," says one Long Beach system leader. As a new leader in any domain, "living up to your word" and getting results quickly are both essential for gaining the trust of stakeholders. Quick wins act like a stake in the ground, emphasizing the seriousness with which the reforms are being undertaken. For example, in order to fulfill its commitment to act against escalating gang violence, Long Beach became the first school district in the U.S. to require students to wear uniforms (so that gang affiliation would not be visible from clothing). It was a fast, highly visible action that made parents feel that there were positive winds of blowing through their system. Similarly, by 2000, three years after Prime Minister Tony Blair's first election, England had already moved from a situation in which just 62 percent of its primary school students achieved its literacy targets to one in which 75 percent did so.

There are times when getting "early wins" on the board as quickly as possible is necessitated by circumstances. The example of Poland in the late 1990s is a case in point. In the fall of 1997, the new government headed by Jerzy Buzek took office following the recent election. In the following spring, Minister of Education Handke announced an education reform program, introducing an extension of general education by one year in order to "give students a chance." The previous school model comprised eight years mandatory primary education and four years secondary, with students split into vocational and academic tracks; roughly fifty percent of all students were in each track. Handke changed the school model, making six years of primary schooling mandatory, followed by three years of general lower secondary education, thereby giving fifty percent of the student population an additional year of general education, relative to the previous system, and three years upper secondary, with students streamed into three tracks, based on their GPA and grade nine exam result. As already described, in consequence of this reform, Poland needed 4,000 lower secondary schools to be open by the start of the new school year. The Ministry gave each municipality a target for the number of schools it needed to provide, leaving it to them to decide whether they would meet this target by ➔

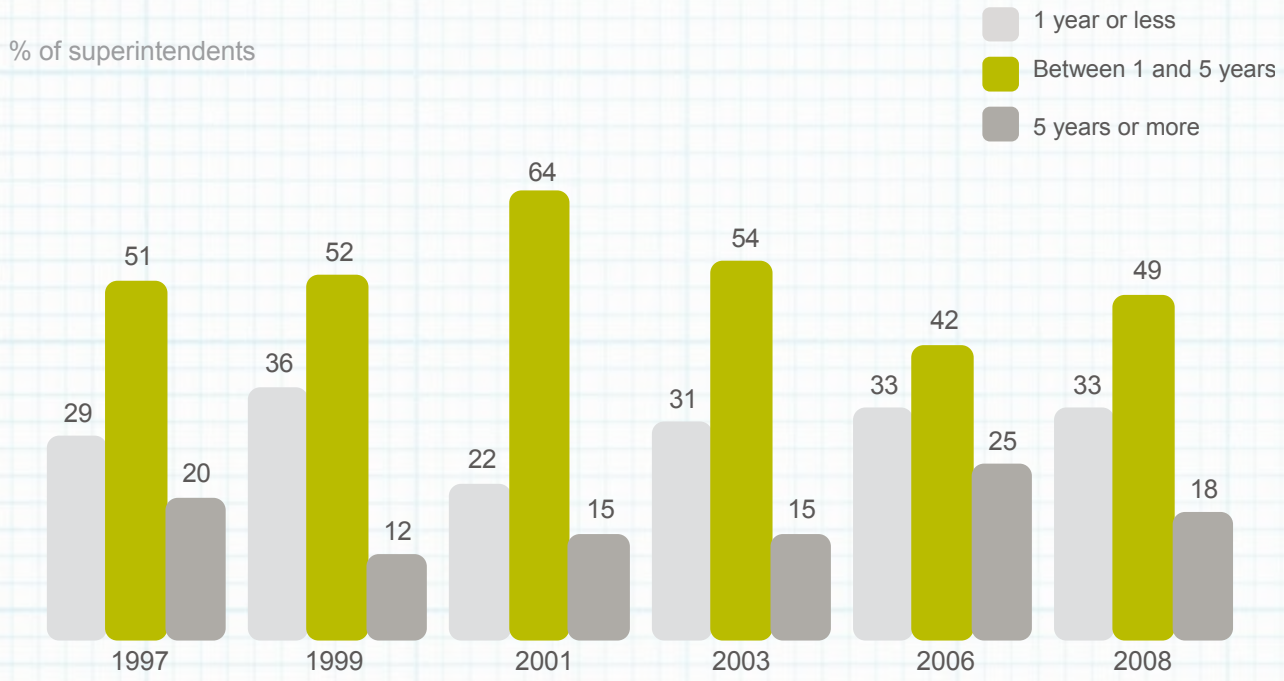
# Exhibit 35:
# US urban superintendants had an average tenure of 2.8 years during 1997-2008




1 The Council of the Great City Schools comprise 66 school districts, considered the largest urban districts in the US. They represent 15% of the nation's K-12 students and 30% of the nation's low-income students, students of color, and English language learners

Source: Council of the Great City Schools[1], Urban Indicator, Winter 2008/2009

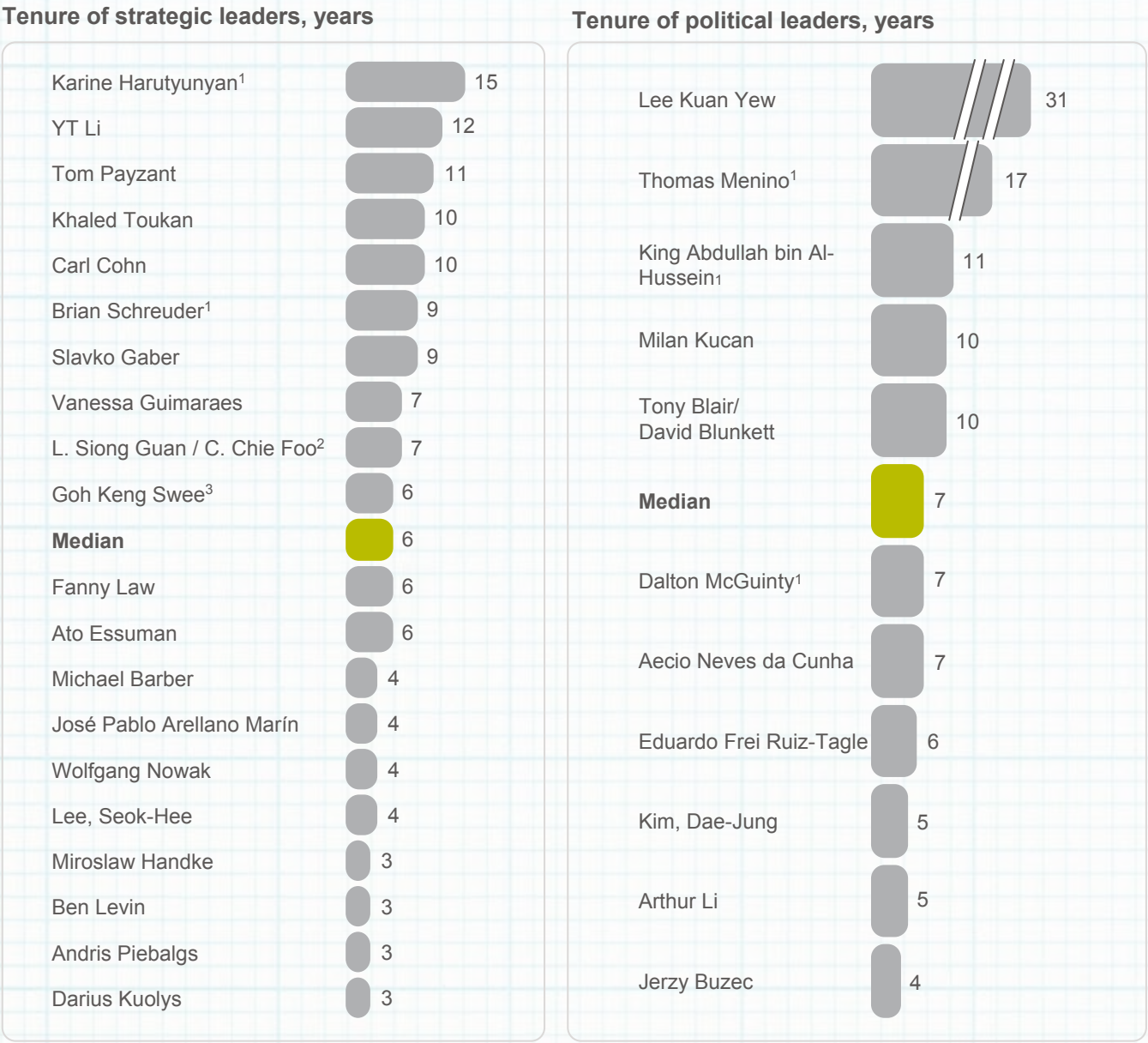
## Exhibit 36:
## Only 18% of urban superintendants had a tenure of five years or more in 2008




1 The Council of the Great City Schools comprise 66 school districts, considered the largest urban districts in the US. They represent 15% of the nation's K-12 students and 30% of the nation's low-income students, students of color, and English language learners

Source: Council of the Great City Schools[1], Urban Indicator, Winter 2008/2009

# Exhibit 37:
# The median tenure of leaders in the systems we studied was six years for strategic leaders and seven for political leaders




1 In position at time of research
2 Lim Siong Guan (Permanent Secretary 1997-1999) and Chiang Chie Foo (2nd Permanent Secretary 1997-1999 and Permanent Secretary 1999-2004) worked closely together to provide strategic leadership for 1997 reforms
3 Although a Minister, Goh Keng Swee was identified by interviewees as having provided the strategic leadership for Singapore's 1980 reforms
SOURCE: System interviews

reconstituting primary schools or by building new lower-secondary schools. For financial reasons, the vast majority of municipalities opted for the former route; this required them to embark on extensive negotiations with their local communities about which primary schools should be reconstituted. By February 28, 1999, all the municipalities had submitted their plans to the Ministry of Education for meeting the target; and, on September 1, 1999, some 4,000 lower secondary schools opened across Poland, ready to receive their first students.

## Staying power

A very important feature of the careers of successful system leaders is that once they are in position, they tend to stay awhile. In the U.S., for example, the urban school district Superintendents in our sample (Long Beach and Boston) had tenures of ten years. By comparison, the overall average tenure of U.S. Superintendents of urban school districts was 3.5 years in 2008, and averaged 2.8 years during the period 1997-2008 (Exhibit 35). Only 18 percent of US urban school Superintendents have been in office for five or more years, while 33 percent have been in office for one year or less (Exhibits 36). Other systems face similar tenure challenges; for example, the average tenure of England's Education Secretaries is just two years, as is that of France's education ministers.

Looking across our 20 improving systems, the median tenure of strategic leaders is six years while that of political leaders is seven years (Exhibit 37). This staying power is critical for leaders to be able to see their reform initiatives through to completion. Says one Long Beach senior staffer, "Continuity lets you have the eye on the prize ... even if it is not perfect the first time you put it out there, you get on a path and can keep refining until you get it exactly how you want it." This point is reinforced by several of our system leaders, who expressed the opinion that their success came late in their tenure; without the time to demonstrate results, their record may have looked quite different today. Tom Payzant, former Superintendent of Boston schools, notes that his revamped math curriculum took four years to show results: "The first three years were flat on performance. But in the fourth year, our student scores went up significantly . . . There had been lots of resistance to the math curriculum, we needed the time and political capital from the Mayor to be able to stick to it and show results." Similarly, Carl Cohn, former Superintendent of Long Beach schools observed: "Our student gains took off in years five to ten. How many School Boards today would give that time to a Superintendent?" Not only did Cohn have a lengthy tenure, but four out of the five School Board members who had hired him stayed in their role for ten years, providing yet another layer of continuity.

For the school system as a whole, of course, what leaders do to ensure continuity in leadership over the longer haul of the improvement journey can be just as important as their personal tenure. We examined how they "build their bench" in some depth in Chapter 3, "Sustaining."

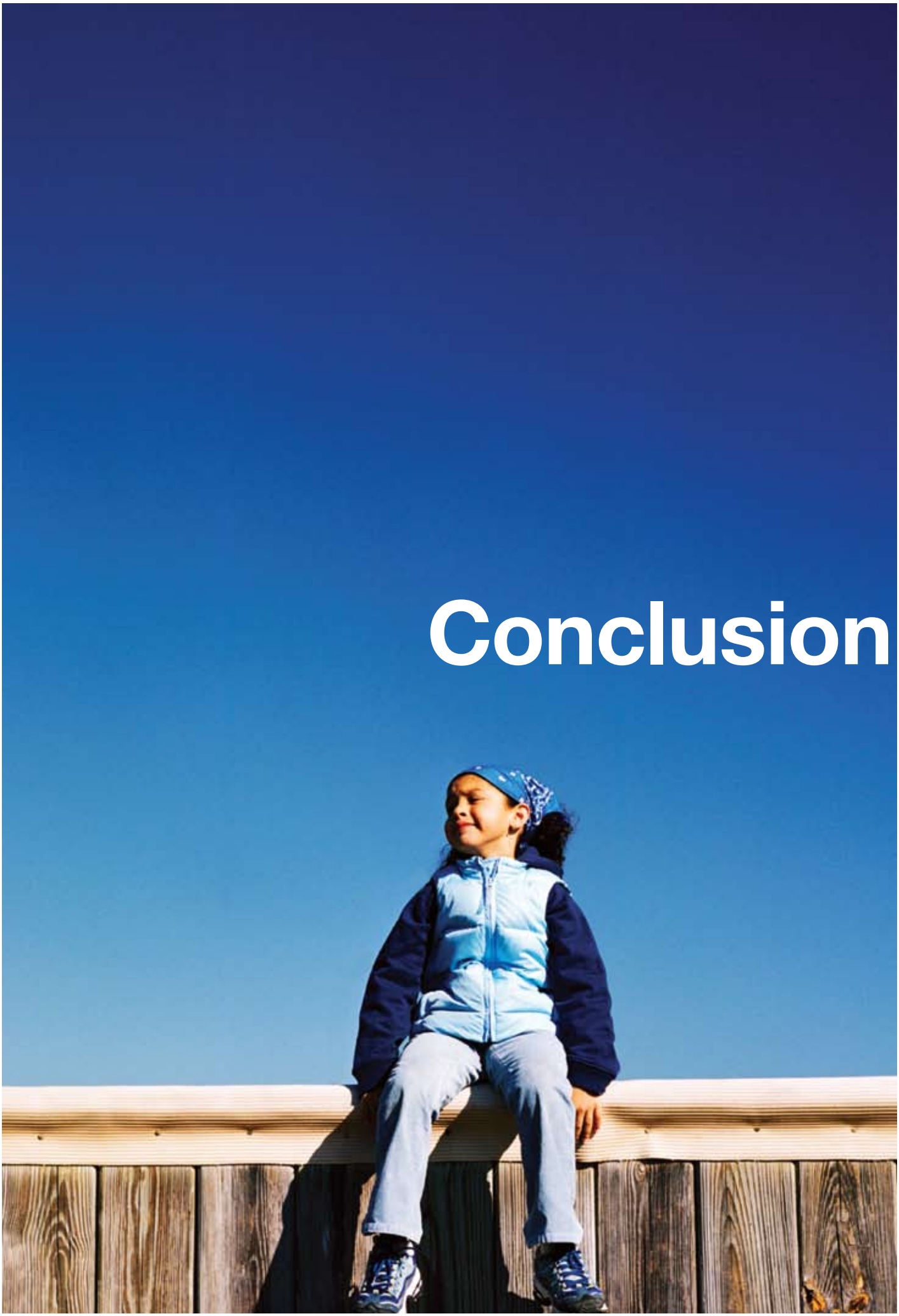
In this chapter we have shown how systems that have successfully ignited schooling reform and then sustained momentum have been able to ignite the reform process by taking advantage of one of three circumstances: a political or economic crisis, or a high-profile report critical of the system's performance, or the appointment of a new political or strategic leader.

Of the three, acquiring a new leader is the most significant factor – all the improving school systems in our sample had appointed a new strategic leader at the start of their reform, and over half had a new political leader, whatever other circumstances prevailed.

Leaders that have been successful in directing their system towards continuous improvement have characteristically taken advantage of the opportunities afforded by the "clean slate" they bring to the role, have followed a common "playbook" of practices, and have benefited from their considerable longevity in the role. Good leadership is critical to success in improving school systems, just as it is in many other fields. ■

ESCOLA

# Conclusion



*To begin with, our knowledge grows in spots. The spots may be large or small, but the knowledge never grows all over: some old knowledge always remains what it was. Your knowledge of pragmatism, let us suppose, is growing now. Later, its growth may involve considerable modification of opinions which you previously held to be true. But such modifications are apt to be gradual.*

– William James, Pragmatism and Common Sense[36]

Nothing is more emotive than education. The quality of our children's schools affects every aspect of their life, shaping the child's personal destiny and the society's capacity for creativity and economic development. This rightly can make school system reform the major focus not just for educationalists but also for political leaders, employers, and parents alike. Often, because of the magnitude of what is at stake in the quality of education provided in our schools, passions run high and debate is heated. What we have tried to do in this report is to disaggregate school system reform into its constituent spots, take a good, long, hard look at the nature of each spot, and then put all the spots back together and examine the overall school system improvement journey from a broader perspective.

We have relied on data drawn from well-respected international authorities that, once standardized, has enabled us to make an objective comparison of different systems. This report records what we have found in looking at each of the spots in the school system improvement journey. It reveals a common pattern that has otherwise largely gone unrecognized; that this is the case is not because other people have not previously seen or appreciated certain aspects of what is reported here – some have done so and, in some cases, in more detail and depth – but largely because the meaning of this overall pattern has been obscured by trees sprouting every which way.

The pattern we have found shows that all the school systems that are successful in achieving sustained improvement within a given performance journey share a common set of characteristics in what they do and how they do it. One reason why this pattern may have been previously obscure could be due to the fact that these characteristics change over time, depending on what stage of the journey the school system has reached. In the early days, outcomes improvement is all about stabilizing the system, reducing variance between classrooms and schools, and ensuring basic standards are met. At this stage of the journey, the reforms are almost always driven from the center. Later, as the system improves, the engine for improvement shifts to instructional practices. This, by its very nature, has much less to do with the centre and is primarily driven by the teachers and the schools themselves: it is all about turning schools into learning organizations. The pattern only becomes clear when this one spot is studied assiduously: without this, it is all too easy to confuse what is needed at one stage with what is necessary at another, quite different, stage.

A second reason why this pattern may have been obscure until now might be that every school system adapts the interventions cluster to its own context. This often makes these interventions appear superficially quite different from one another, disguising their commonality. Only by disaggregating the exact content of each intervention has it been possible to identify the nature of the wood. A tree might be a tree, but to Linnaeus, this is entirely to miss the point. Only when the architecture of families is first understood is the naming of parts truly meaningful.

The Lewis and Clark expedition provides a corollary. During the early years of the nineteenth century, following the Louisiana Purchase, Lewis and Clark led the first expedition across the United States' new territory. Their goal was to establish the lay of the land and to gain an understanding of it. As they journeyed across the new terrain of the North West, they collected hundreds of botanical, zoological and mineral specimens; they documented the extent of the Rocky Mountains and the channels and sources of the great rivers; they described the places they went and the people they came across. It was Lewis and Clark who produced the first meaningful maps of the United States, including the sources of the Missouri and Columbia rivers and their relationship to the Rocky Mountains. Their expedition forms a benchmark for all future understanding of the territory. They named the parts, forming the basis upon which others built.

This report seeks to be a map of school systems' performance journeys. It identifies the intervention clusters, and locates the importance of contextualizing, sustaining, and ignition on the journey. This leaves much that is yet unknown and still to be explored. We hope this report will encourage school system leaders and education researchers to develop further knowledge ➔

about improving school systems, filling in some of the blanks left on the map. Here are just a few suggestions about potential areas of further study:

- How do the improvement journeys and outcomes of systems with similar context vary? For example, this could be answered by studying 'matched pairs' such as two states in the same country.

- Where is the line drawn in contextualizing interventions? Is there a point at which a system compromises the intervention cluster by excessively contextualizing the interventions?

- What are the prerequisites and sequence of the interventions within each cluster? Do patterns exist that are more likely to be successful than others?

- How do schools systems successfully devolve authority and direction to the middle layer and to the schools? How do systems develop the skills of the middle layer?

- What is missing from the map? For example, there are as yet blanks in the map of the elements and development of collaborative practice.

None of this is to suggest that school system improvement is either science or art: it is neither. It is the disciplined craft of repeated practice and learning within the context of the system: the practice and internalization of the pedagogy. This practice requires institutional support and is one reason why Peter Senge objects to viewing teachers as practicing in isolation. It is not about the individual's skills, but a skilled system: "The traditional approach to helping educators has been to develop the skills of individuals to do their work better [as opposed to] enhancing the collective capacity of people to create and pursue overall visions."[37] For the improving schools in this study, these visions are about continuing improvement.

The school systems examined in this report show that the improvement journey can never be over. Achieving and sustaining a school system's progress is very hard work, and systems must keep expending energy in order to continue to move forward: without doing so, the system can fall back, and thereby threaten our children's well-being. Our hope is that this report has provided an overview of the school system improvement journey, and has given the first outlines of the landscape that will be navigated further by education's future explorers. ■

# Appendix

# I. School system selection

## A. Criteria for system selection

In selecting school systems, we sought to meet two objectives: 1) to select systems that have achieved clear improvement in student outcomes, and 2) to compile a diverse sample of systems so that we could learn what was unique versus what was universal, thereby ensuring our insights have wide-ranging relevance. We defined diversity along several dimensions – by size of system, location (representing five continents), starting performance levels (on student assessments), and system type (centralized and decentralized; private and public systems).

In selecting this diverse set of improved systems, we established two sets of criteria. The first enabled us to identify "Sustained Improvers": systems that have a long history of reform and consistently see improvement. The second set of criteria enabled us to identify "Promising Starts": systems that have only recently begun reform efforts, but which have seen significant improvement in a short period of time. Promising Starts are restricted to systems in developing countries and emerging markets that despite not having a long history of international testing, have shown remarkable improvement in the assessments in which they have participated, and embody an improvement journey that has employed innovative techniques and strategies. A list of systems in both categories appears in Exhibit 38.

Whether a system is classified as a Sustained Improver or a Promising Start, it has had to demonstrate significant, sustained, and widespread improvement to be included in this study. We used several international assessments to establish whether a system qualifies in either category or not:

- 1964 First International Mathematics Study (FIMS)
- 1970 First International Science Study (FISS)
- 1978 First International Mathematics Study (SIMS)
- 1983 First International Science Study (SIMS)
- 1995, 1999, 2003, and 2007 Trends in International Mathematics and Science Study (TIMSS)
- 2000, 2003, 2006 Program for International Student Assessment (PISA)
- 2001, 2006 Progress in International Reading Literacy Study (PIRLS)
- We also used national and state/regional assessments for school systems that do not participate in international assessments:
- 1971–2009 National Assessment of Educational Progress (NAEP) for US school systems
- 2005, 2007, 2009 Index of Development of the Basic Education (IDEB) for Minas Gerais, Brazil
- 2006, 2007, 2008 Annual Status of Education Report (ASER) for Madhya Pradesh, India
- 2000–10 California Academic Performance Index for KIPP, Aspire

For most of our systems, the 1995 TIMSS assessment is the earliest source of student performance data; the 1995 assessment was the first occasion when TIMSS used the 500-scale that created consistency in the distribution of scores over time and thereby helps the comparison of results over time. For all countries participating in 1995 TIMSS, the mean score was adjusted to 500 with a standard deviation to 100. All subsequent TIMSS exam data was also placed on this metric, thereby enabling comparison between the scores of countries across the different years of the TIMSS tests.

## B. Criteria for sustained improvers

Sustained Improvers exhibit Significant Gains, Sustained Gains, and Widespread Gains.

### Significant Gains

Systems that qualify as Sustained Improves with Significant Gains are divided into three main groups: ➔

# Exhibit 38:
# Our school system sample comprises sustained improvers and promising starts

**Sustained improvers**

Systems that have sustained improvement with 3 of more data points over 5 or more years

**Promising starts:**

Systems that have started improving as represented by ongoing improvement with just 2 data points or less than five years of improvement

| Systems | Time period of assessment | Sustained improvers | Promising starts |
|---|---|---|---|
| 1. Singapore | 1983 – 2007 | ✓ | |
| 2. Hong Kong | 1983 – 2007 | ✓ | |
| 3. South Korea | 1983 – 2007 | ✓ | |
| 4. Ontario, Canada | 2003 – 2009 | ✓ | |
| 5. Saxony, Germany | 2000 – 2006 | ✓ | |
| 6. England | 1995 – 2007 | ✓ | |
| 7. Latvia | 1995 – 2007 | ✓ | |
| 8. Lithuania | 1995 – 2007 | ✓ | |
| 9. Slovenia | 1995 – 2007 | ✓ | |
| 10. Poland | 2000 – 2006 | ✓ | |
| 11. Aspire Public Schools, USA | 2002 – 2008 | ✓ | |
| 12. Long Beach, CA, USA | 2002 – 2009 | ✓ | |
| 13. Boston/Massachusetts, USA[1] | 2003 – 2009 | ✓ | |
| 14. Armenia | 2003 – 2007 | | ✓ |
| 15. Western Cape, South Africa | 2003 – 2007 | | ✓ |
| 16. Chile | 2001 – 2006 | | ✓ |
| 17. Minas Gerais, Brazil | 2003 – 2008 | | ✓ |
| 18. Madhya Pradesh, India | 2006 – 2008 | | ✓ |
| 19. Ghana | 2003 – 2007 | | ✓ |
| 20. Jordan | 1999 – 2007 | | ✓ |

1 Primary focus was on Boston, within the context on Massachusetts State Reforms. Massachusetts NAEP results also indicate the state as a sustained improver from 1998-2007 on mathematics and reading

Source: McKinsey & Company interventions database

- Systems exhibiting improvement prior to the 1995 TIMSS: namely Hong Kong, Korea, and Singapore. The criteria for systems whose improvement journey started prior to the 1995 TIMSS' assessment is that these systems had to be ranked in top-five school systems on PISA (2000) or TIMSS (1995) on their first assessment and that there should be clear evidence of a clear improvement trend prior to the 1995 assessments. The three systems topped international assessments in their first testing on TIMSS and PISA. We therefore used the earliest available international assessment data to provide evidence on system improvement prior to 1995, namely FIMS, FISS, SIMS, and SISS, in order to analyze the full improvement journey of these systems.

- Systems exhibiting improvement from 1995 onwards and that participated in international assessments: such systems need to demonstrate an improvement greater than or equivalent to 25 percent of a school-year equivalent on PISA or TIMSS assessments. A gain on PIRLS is considered as reinforcing evidence.

- Systems that have not participated in international assessments: these systems need to demonstrate the following criteria to qualify as Sustained Improvers: states/provinces have significantly outpaced the average on national assessments; districts that have significantly outpaced the average on state/provincial assessments; school networks have outpaced the districts in which they operate on state assessments.

### Sustained Gains

- Sustained Gains for Sustained Improvers: this is defined as the system having achieved five years or more of improvement, with at least three data sets indicating an upward trend.

### Widespread Gains

- Widespread Gains for Sustained Improvers: is defined as gains demonstrated across multiple subjects and/or assessments. Reducing variance (e.g. between school variance on PISA) is considered to be reinforcing criteria for selection.

## C. Criteria for promising starts

Promising Starts, similarly to Sustained Improvers, exhibit Significant Gains, Sustained Gains, and Widespread Gains.

### Significant Gains

Promising Starts that exhibit Significant Gains are of two types:

- Systems that participated in international assessments: such systems need to demonstrate an improvement greater than or equivalent to 25 percent of a school-year equivalent on PISA or TIMSS assessments. A gain on PIRLS is considered to be reinforcing criteria.

- Sub-systems (regions/states, school networks, etc) that did not participate in international assessments: these systems need to demonstrate that they have significantly outpaced their national or regional average over time on an objective and consistent student outcome metric.

### Sustained Gains

Sustained Gains made by Promising Starts: is defined as systems that have achieved at least two to three years of improvement, with at least two data sets indicating an upward trend.

### Widespread Gains

Widespread Gains made by Promising Starts: requires that gains be made in at least one high-priority area (science, math, literacy) provided that all performance data in that area shows consistent gains. Reducing variance (e.g. between the highest and lowest- performing students) is considered to be reinforcing criteria for selection.

## II. The universal scale

One of the critical underpinnings of our work has been producing a Universal Scale by which we are able to classify school systems' performance as poor, fair, good, or great. The systems in our selection participated in various assessments (TIMSS, PISA, PIRLS, NAEP) across multiple subjects (math, science, reading), at a variety of grades/levels (primary and lower secondary) and over a prolonged period, with test dates from 1995 until 2007.

Collectively, there were 25 unique assessments, each using an independent scale.

## A. Systems participating in international assessments and NAEP

We used the methodology of Hanushek et al.38 to normalize the different assessment scales of the systems in our selection that have participated in international assessments or NAEP into a single Universal Scale. The units of the Universal Scale are equivalent to those of the 2000 PISA exam; on this scale 38 points is approximately equivalent to one school year. For example, eighth graders in a system with a Universal Scale score of 505 would be on average two years ahead of eighth graders in a system with a Universal Scale score of 425.

To create the Universal Scale, the Hanushek methodology requires calibrating the variance within individual assessments (e.g. PISA 2000) and across every subject and age-group combination; this was done for the 39 different assessments relevant to our sample systems dating back to 198039. There are numerous challenges in calibrating variance. Each of these assessments tests different school systems, reflecting multiple geographies, wealth levels and demographics. For example, PISA predominately includes OECD and partner countries while TIMSS has a much larger representation that includes developing nations. A variance of X on TIMSS is therefore not equivalent to a variance of X on PISA. Within each assessment, the cohort of participating countries changes from one year to the next. In order to compare the variance between the two assessments, a subset of mature and stable systems (i.e. those with consistently high rates of school enrolment) is used as a control group, and the variance between these systems is then compared across the assessments. After calibrating the variance, the methodology calls for calibrating the mean for each assessment. This has been done using the U.S. NAEP assessment as a reference point. The U.S. NAEP was selected for this purpose firstly because it provides comparable assessment scores as far back as 1971 and secondly because the U.S. has participated in all international assessments.

Once the various assessment scales have been made comparable, each school system's average score for a given assessment year is calculated by taking the average score across the tests, subjects, and grade levels for that year. This creates a composite system score on the universal scale for each year that can be compared over time.

Finally, each country's Universal Scale score is classified either as poor, fair, good, great, or excellent, based on the distribution below. None of the systems in our sample exceeded the threshold requirement for Excellent. The various performance categories are explained below:

- Excellent: greater than two standard deviations above the mean
- Great: greater than one standard deviation above the mean
- Good: less than one standard deviation above the mean
- Fair: less than one standard deviation below the mean
- Poor: greater than one standard deviation below the mean
- Exhibit 2: Illustrative distribution of the Universal Scale scores

According to the distribution of scores on the Universal Scale, the improvement gap – the improvement required for a system to progress from one performance level to the next – is 1 school-year equivalent, or 38 Universal Scale points. The baseline score are as follows: Excellent > 560 points; Great 520-560 points; Good 480-520 points; Fair 440-480 points; and, Poor <440 points (see Exhibit 39).

## B. Systems not participating in international assessments and NAEP

For the school systems in our selection that have not participated in international assessments or NAEP, including Aspire Public Schools, Madhya Pradesh, and Minas Gerais, we approximated their position on the Universal Scale using available data.

### Aspire Public Schools

We placed Aspire Public Schools' journey on the ➔

Universal Scale at first as "fair" and then moving to "good." In order to reach this conclusion, we derived Aspire's equivalent average score on NAEP by calibrating the California API score of Aspire schools against the scores of Los Angeles Unified School District and California state, which report both API and NAEP scores. The NAEP score was then normalized to the Universal Scale as per the Hanushek methodology, described above. There is also significant evidence to show that Aspire is improving significantly and at a much faster pace then peer school districts in California. For example, Aspire schools in both Oakland and Stockton school districts not only outperform the average for schools in their districts, but also "out-improve" them (i.e. improve at a faster rate than other schools); their improvement rate is over 200 percent the average in Oakland and over 400 percent that in Stockton.

### Madhya Pradesh

We placed Madhya Pradesh on the universal scale as "poor." Although no assessments exist which could directly link Madhya Pradesh's performance to international assessments, the evidence shows that it is struggling with basic literacy and numeracy in primary and secondary students: therefore, by qualitative measures it is significantly lower performing than fair systems, such as Chile and Armenia. The case for improvement in Madhya Pradesh is also strong, given evidence from ASER. As seen in Exhibit 6, Madhya Pradesh shows improvement in student outcomes in both mathematics and reading. In comparison, during this period the Indian average score has declined in mathematics and stagnated in reading.

### Minas Gerais

We placed Minas Gerais on the universal scale as "poor." Although no tests exist which could directly link Minas Gerais' performance to international assessments, Brazil's overall performance on TIMSS is well below the cut-off for "fair." The most recent Brazilian IDEB results show evidence of Minas Gerais' significant improvement, as it has moved from being ranked fifth in the nation in 2005 to being its highest performer in 2009. Furthermore, the provincial Proalfa assessments in Minas Gerais show a 76 percent improvement in those achieving the recommended reading proficiency level between 2006 and 2010.

## III. School System Interventions Database

During the research for this report we visited each of the school systems to learn of their experiences, conducting interviews with approximately 200 system leaders and their staff who had implemented reforms. This gave us a good understanding of all the improvement interventions that had been made in each system. We created a database to capture the nature of each reform intervention taken by the school system and when it was taken.40

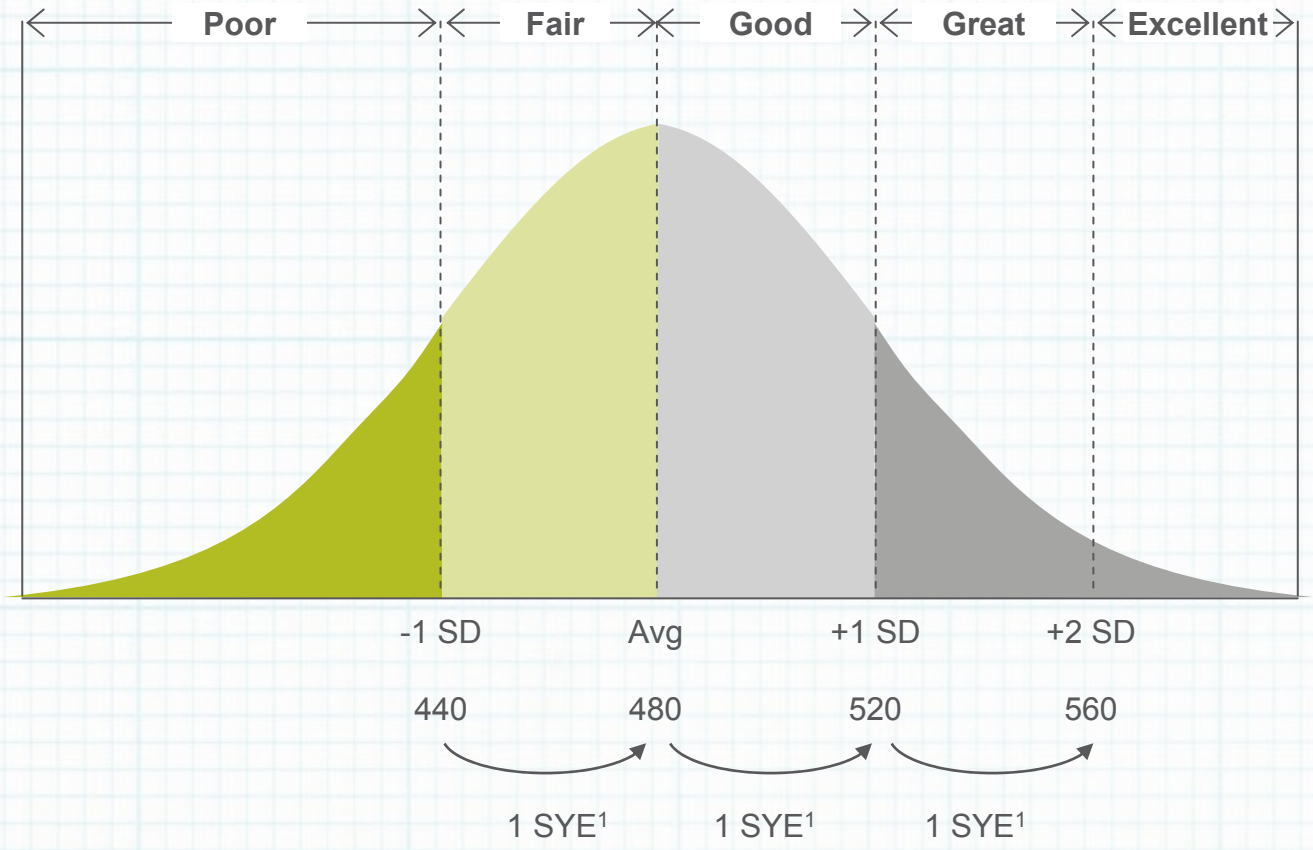
We used the information gleaned from the interviews, in combination with a review of the official literature on the education reform, to capture all the interventions that were identified as significant – either those that "made a difference" or that had required heavy investment of human or financial resources. For example, we recorded reforms that led to the creation of new academic achievement standards for primary students, or the installation of a quality assurance board that assessed school performance. At the end of each interview, we added the 15-30 most relevant and highest impact interventions to the database.[40]

By cross-referencing the information on this database with other system indicators, such as size, performance level, geography, governance type, we are able to explore a wide range of dimensions: e.g. the types of interventions used in "poor" systems as compared to those used in "fair" systems, or the interventions that are unique to large school systems.

### A. Structure of the interventions database

To create this database, we mapped each reform intervention taken by systems onto the period of their history when the reform was taken. Overall, the database identified approximately 575 individual interventions taken by school systems as part of their school system reforms. We identified approximately 60 unique intervention types, such as teacher training, school vouchers, increasing teacher salaries, and redesigning the curriculum. Then, each time an intervention occurred in a system, we catalogued it along with the ➔

## Exhibit 39:
## Illustrative distribution of the universal scale scores




1 School year equivalent

Source: McKinsey & Company

unique specifications of the reform intervention. For example, while both Madhya Pradesh and Long Beach were classified as having used "teacher coaches", the specifications of how each system uses teacher coaches were different. In addition, each intervention was also pegged to a "recipient" – i.e. the stakeholder, including students, teachers, principals, parents, the center, or the entire school system – that was directly influenced by the intervention.

To help us reveal the trends in the interventions used, we overlaid a framework of areas and sub-areas onto the 575 individual interventions. For example, in the area of "accountability", we identified three sub-areas: performance assessment, inspections, and self-evaluation. This enables us to produce a number of different views of the data. For instance, when we aggregate this data, we can establish how many times our improving systems have utilized "accountability" as an intervention overall, or looking at a more specific view, how many systems have used "teacher evaluations".

Finally, we also mapped each intervention onto a framework of intervention types that identifies whether the intervention is a process, structure, or resource intervention. This allows us to explore, for example, how often and when systems have injected resources compared to how often and when they have changed a process.

## B. Analyses emerging from the database

While the interventions database could potentially be used to run a myriad of different analyses by cross-referencing any information about the school system with another, or by overlaying any framework onto the interventions, this report has primarily focused on two analyses that have guided and directed our thinking:

- The frequency of application of each intervention or intervention area
- The clusters of interventions used at each performance level

### Frequency of application of each intervention or intervention area

The simplest analysis we performed was to determine how often an intervention occurred. This analysis enabled us to answer a number of important questions: for instance, how often a school system offered "support" to teachers through such interventions as teacher coaches, increased salaries, and increased training. We could also answer questions such as, "Do improving school systems employ a greater portion of support interventions versus accountability interventions?" and, "Do school systems use accountability interventions more frequently with teachers or with principals?"

### Clusters of interventions at each performance level

The primary objective of the cluster analysis was to identify what suite or cluster of reform interventions our sample systems employed as they moved from one performance level to another. Our starting point for this analysis was a prevalence calculation: "What percent of the systems in each performance stage employed a particular intervention?"

Our second step was to answer the question, "How distinctively is each intervention associated with, or weighted to, each performance stage?" For example, an intervention that occurs in 40 percent of all reforms journeys is evenly weighted (or not distinctively associated with one stage), whereas another intervention that occurs in 25 percent of fair to good journeys and 5 percent of each of the other stages, is weighted towards the fair to good improvement journey. This is true even though the prevalence of the first intervention in fair to good (40 percent) is higher than in the second intervention (25 percent). Therefore, to carry out this analysis we first needed to normalize our prevalence calculations by setting the sum of each intervention's prevalence to a common scale (e.g. 100); this enabled us to make comparisons across interventions within a given performance journey. To illustrate this, in our hypothetical example, in which there is a 25 percent prevalence in the fair to good journey and 5 percent in each of the other stages (5-25-5-5) versus 40 percent overall

in all four stages (40-40-40-40), the normalized values would be (13-63-13-13) and (25-25-25-25) respectively.

However, that still leaves the question of, "How remarkable is the weighting of each initiative in the specific performance journey to which it is most strongly linked?" To establish this, we calculated the distribution (mean and standard deviation) of the normalized weightings of all the interventions within each performance stage. In our hypothetical example, this distribution analysis would tell us – when looking across all the interventions in fair-to-good stage – whether the weighting of 63 in the fair to good journey is relatively high or not. This analysis enables us to identify the set of interventions that are most distinctively associated with each performance journey stage.

The final question we asked in our quantitative analysis was, "What degree of correlation is there between the various interventions in each improvement stage?" – that is, how often do they occur together? In order to answer this question we determined the co-occurrence of interventions within each system's reform journey.

Lastly, we filtered our results to ensure consistency using a qualitative test. There were two parts to this test. First we weeded out true anomalies, such as interventions that only occur once or twice and therefore indicate a very high weighting but cannot reasonably be considered to be part of a universal cluster of interventions. Second, we checked each of the intervention clusters for their resonance with the experiences of each of the system reforms based on our understanding gained from the interviews with that system's leaders.■

# Footnotes

1 Only Finland has so far reached "excellent" globally, though several systems studied here are well advanced along the journey towards it.
2 We took the starting point of the reform as defined by the system leaders themselves, and began our data gathering from this point. Therefore, our database does not capture interventions which pre-date this starting point and which may have influenced the reform journey.
3 These assessments include TIMSS, PISA, PIRLS, NAEP.
4 Hanushek, et al., "The High Cost of Low Educational Performance," OECD, 2010.
5 McKinsey & Company, How the World's Best-Performing School Systems Have Come Out on Top, 2007.
6 Joel Klein has used the phrase "You can mandate awful to adequate but you cannot mandate greatness; it has to be unleashed."
7 Shulman, Lee S., "Signature Pedagogies in the Professions", Daedalus, Summer 2005.
8 Michael Fullan has written extensively on collaborative capacity, which he defines as "Collective capacity is when groups get better—school cultures, district cultures and government cultures. The big collective capacity and the one that ultimately counts is when they get better conjointly—collective, collaborative capacity, if you like. Collective capacity generates the emotional commitment and the technical expertise that no amount of individual capacity working alone can come close to matching..." All Systems Go, 2010.
9 We use the term 'strategic leader' to refer to the individual responsible for the strategic direction and implementation of the school system reform
10 We took the starting point of the reform as defined by the system leaders themselves, and began our data gathering from this point. Therefore, our database does not capture interventions which pre-date this starting point and which may have influenced the reform journey.
11 The universal scale is based on the Hanushek methodology; see Appendix for full description.
12 Ghana had substantial gains on TIMSS Science (greater than one school-year equivalent) and Math (75 percent of a school year equivalent) for 8th graders from 2003-07, although starting from a low base. Chile had significant gains in both PISA reading (75 percent of one school-year equivalent) and science (50 percent of one school-year equivalent) in 2000-06. While also starting from a low base, Chile had the highest gains of any system on PISA science since 2000.
13 The Department of Education's central team trained 1,500 regional trainers who then, in turn, trained a further 4,500 instructional specialists who then trained 10,000 teachers.
14 The eight areas were: [1] teacher support and development; [2] changes to classroom practice; [3] pre-school/ECD strategy; [4] Advocacy, Family and Community literacy; [5] research; [6] monitoring and support; [7] Coordination and sustainability; and [8] Learning and Teaching Support Material.
15 According to the Northwest Evaluation Association (NWEA) rating of the standards of assessments in the United States.
16 In 2004-05, the program was piloted in four schools. In 2005-06, the pilot was extended to 15 schools, and from 2006-07 it was rolled out to all 47 primary schools across the district.
17 Michael Fullan refers to this phenomena as 'lateral learning', comprising three change forces: 1) mutual allegiance on a large scale (educators' sense of identity gets enlarged and they start to identify with broader system peers); 2) collaborative competition (educators try to outdo themselves and each other); and, 3) development of a shared vision. All Systems Go, 2010.
18 Decentralization of pedagogical rights refers to districts/schools/teachers being given more control over curriculum, setting standards, and/or defining a new instructional program. In some cases, the center enlists the help of schools in desigining new content (as in Armenia, Latvia, and Lithuania). In other cases, the center asks districts / local boards to be responsible and accountable for the learning outcomes of schools (as happened in Ontario in 2009).
19 Initially these school clusters were 5-7 schools, but later were expanded to 12-14 schools.
20 Bonus pool varies each year.
21 A significant share of the training sessions occur within their own schools and classrooms.
22 Order of the Ministry of Education of the Republic of Lithuania of 19 February 1998 No. 331 on Teacher Credentialing Regulations.

23 In the Hong Kong context, 'strongly recommend' is interpreted as 'schools have to do', but some flexibility exists to allow schools to exercise administrative and professional judgement. In the case of professional development, principals would normally give guidance to teachers that they are required to comply.
24 ACTEQ (Advisory Committee on Teacher Education and Qualifications), 2003
25 Hong Kong Education Department, August 1997. "Medium of Instruction, Guidance for Secondary Schools."
26 Tsang, W. 2004. Evaluating the Medium of Instruction Policy in a Post-Colonial Society: The Case of Hong Kong Special Administration Region
27 Ministry of Education. 1969. Education in Singapore.
28 Lee Kuan Yew, 2000
29 Hong Kong adopted a publicly funded, privately operated system in order to be able to resource the rapid expansion of its schools. Under this arrangement, private entities provide 20 percent of the capital expenditure for schools, while the government provides the remaining 80 percent of capital expenditure and 100 percent of operating expenses. Since the 1960s, the vast majority of Hong Kong's schools have functioned in this manner; currently 90 percent are part of this arrangement.
30 Shulman, Lee S., "Signature Pedagogies in the Professions", Daedalus, Summer 2005.
31 Hattie, John. Visible Learning, Routledge, London: 2009.
32 Fullan, Michael. All Systems Go. 2010
33 Interview with Mary Jean Gallagher, Chief Student Achievement Officer, Ontario. November 2009.
34 Each subsequent level came with a ten percent pay increase, which was above the normal yearly rise.
35 We use the term 'strategic leader' to refer to the individual responsible for the strategic direction and implementation of the school system reform
36 James, William (1907) "Pragmatism and Common Sense". Lecture 5 in Pragmatism: A new name for some old ways of thinking. New York: Longman Green and Co: 63-75.
37 Senge, P. (1995) in O'Neil, J. "On Schools as Learning Organisations: A conversation with Peter Senge" Education Leadership Vol.52 No.7.
38 Hanushek, Eric and Ludger Woessmann. "The High Cost of Low Educational Performance: The Long-Run Impact of Improving PISA Outcomes", Paris: Organization for Economic Cooperation and Development, 2010.
39 1 IEA (SISS), 4 TIMSS, 3 PISA, 2 PIRLS, 19 NAEP, and 10 CA API.
40 We took the starting point of the reform as defined by the system leaders themselves, and began our data gathering from this point. Therefore, our database does not capture interventions which pre-date this starting point and which may have influenced the reform journey.

# Contact

education@mckinsey.com

Education
November 2010
Designed by Media & Design | London
Copyright © McKinsey & Company
www.mckinsey.com