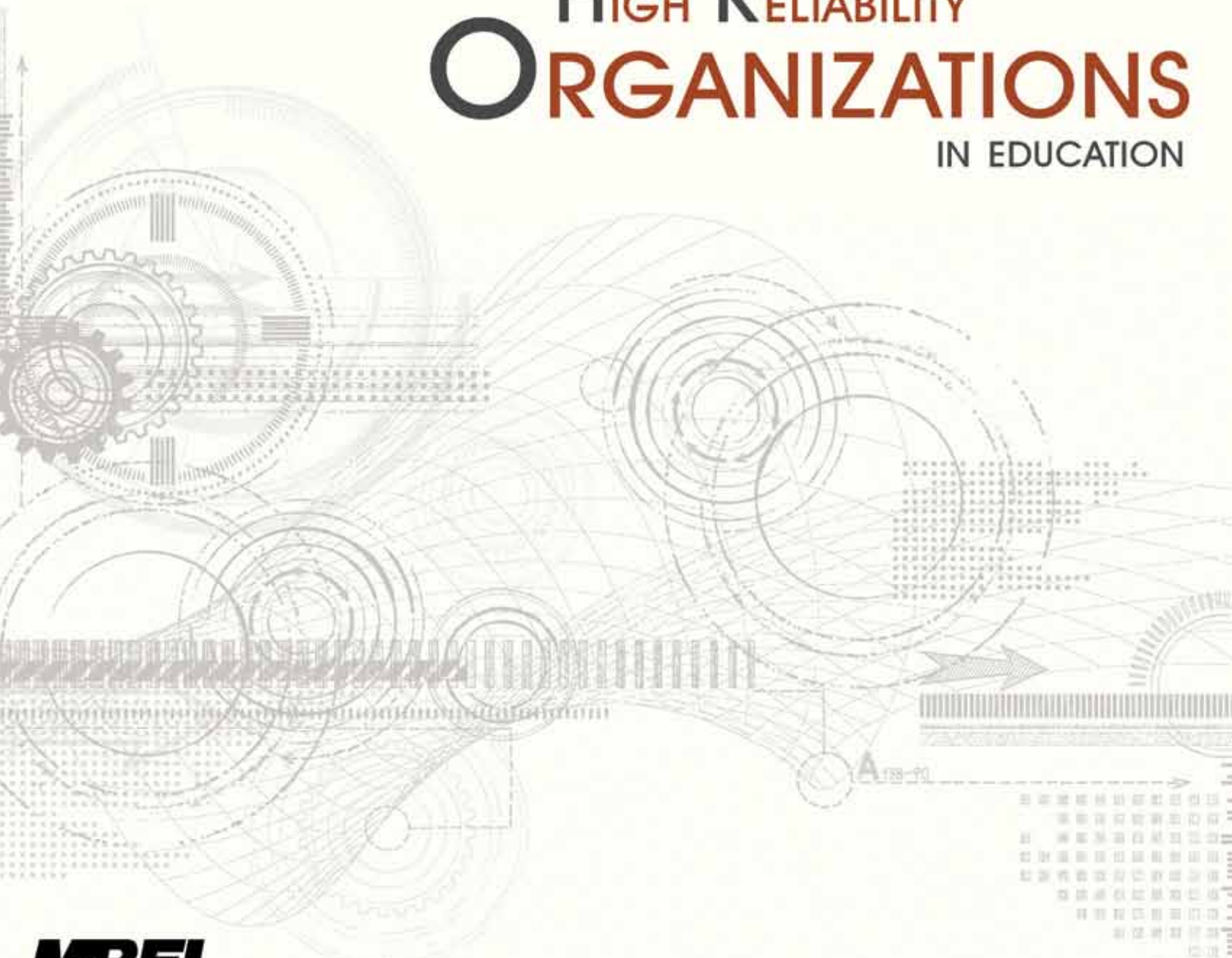
Noteworthy
PERSPECTIVES

High Reliability Organizations in Education



McREL

Mid-continent Research for Education and Learning (McREL) is a private, 501 (c)(3) education research and development corporation. At McREL's headquarters in Denver, Colorado, and McREL's Pacific Center for Changing the Odds in Honolulu, Hawaii, you'll find experienced researchers and education consultants working together to provide educators with research-based, practical guidance on the issues and challenges facing K–16 education.

In January 2011, McREL formed The Network for Innovative Education™ as a call to action for educational systems in the United States to once again become the "best in the world." We feel an urgency to call attention to the "other achievement gap" in our country—the one between our highest performing systems, schools, and students, and those in other countries. This *Noteworthy* tells the story of our beginning work.

Contact McREL to order additional copies of
Noteworthy Perspectives: High Reliability Organizations in Education at:

Mid-continent Research for Education and Learning
4601 DTC Blvd., Suite 500
Denver, CO 80237
P: 303.337.0990 • 800.781.0156 • F: 303.337.3005
E-mail: info@mcrel.org • website: www.mcrel.org

Graphic design by Natalie Voltes

20110906

Noteworthy
PERSPECTIVES

High Reliability Organizations in Education

McREL

Table of Contents

Chapter One: Introduction: Becoming the "Best in the World "at Educating Our Students 1

Chapter Two: Toward Highly Reliable, High-Quality Public Schooling 6

Chapter Three: High Reliability & Leadership for Educational Change 24

Chapter Four: Best in the World: High Performance with High Reliability 36

Chapter One: Introduction

Becoming the Best in the World at Educating Our Students

By James H. Eck, McREL

Once at the forefront of educational achievement by any number of measures, the United States has dropped in standing relative to its economic partners and competitors in the global marketplace. While we have taken incremental steps to improve student learning, other countries have made tremendous strides, not just catching up to us, but passing us by. As a result, U.S. students are now grossly under-represented in the ranks of top scorers on international tests, putting America's ability to maintain its position as a leader in innovation, technology, and prosperity at risk.

The Programme for International Student Assessment (PISA) exam is a battery of assessments of 15-year-old students in mathematics, science, and reading, administered internationally every three years. In 2006, the United States ranked 25th of 30 nations in mathematics and 24th of 30 in science (reading literacy was on an off-year). This puts us on par with Spain, Portugal, and the Slovak Republic, and far behind Canada, the Netherlands, Australia, Korea, and other countries. This, in spite of the pledge of the National Governors Association in 1989 that U.S. students would lead the world in mathematics and science achievement by 2000 (Walberg, 2003).

The United States has participated in all of the international tests since 1964 and has maintained a longitudinal testing system of its own, the National Assessment of Educational Progress (NAEP). Stanford economist Eric Hanushek and his colleagues (Hanushek, Jamison, Jamison, & Woessmann, 2008) conducted a cross-assessment analysis of a variety of international tests in reading, mathematics, and science administered between 1964 and 2003 and calibrated each of the separate international tests to the NAEP. They found that performance of U.S. students over the past four decades remained flat (along with Germany and Hungary), while those of students in the Netherlands and Finland have improved. Asian countries have consistently performed well, including those that have entered the international assessment arena along the way. Poland was the most rapidly improving nation, registering average gains on the PISA of more than 25 points between 2000 and 2006 alone and surpassing the U.S. performance in 2006. (Hanushek, et al., 2008; McKinsey & Company, 2009; Organisation for Economic Co-operation and Development [OECD], 2010).

A series of research studies has led Mid-continent Research for Education and Learning (McREL) to an important conclusion about how to improve the international competitiveness of U.S. schools. As in the field of medicine, where some advances in ensuring better end results for patients have come not from technology or pharmaceuticals, but from simple improvements, including better hygiene in hospitals and the systematic use of checklists, what we most need now in U.S. education is not more funding, more technology, or radical new innovations. Rather, the solution to ensuring America's students are able to compete on a global stage lies in re-thinking the role of schools and school districts. By assisting school systems to more closely resemble "high reliability" organizations (HROs) that already exist in other industries and benchmarking against top-performing education systems from around the globe, America's school systems can transform themselves from compliance-driven bureaucracies to world-class organizations.

Top systems in the world

- Alberta, Canada
- Australia
- Belgium
- Finland
- Hong Kong
- Japan
- Netherlands
- New Zealand
- Ontario
- Singapore
- South Korea

Missing the bar

Research carried out by consultancy firm McKinsey & Company between May 2006 and March 2007 resulted in the report, *How the World's Best-*

Performing School Systems Come Out On Top (2007), in which the authors identify the highest performing school systems by their 2006 performance on the PISA exam. In addition, the McKinsey researchers had conducted extensive site visitations and included a comprehensive review of current literature, and interviews with experts, policymakers, and practitioners.

Blatantly missing from this list of top performers are school districts from the United States, although several (Atlanta, Boston, Chicago, and New York City, along with one state—Ohio) were identified as being on "strong improvement trajectories." From an economic standpoint, which was one impetus for conducting such a study in the first place, the results are discouraging. Despite huge increases in spending for education and ambitious reform efforts, Americans are seeing little improvement across their school systems. Few of the most widely supported reform strategies (e.g., giving schools more autonomy, reducing class sizes) have produced the promised results (McKinsey & Company, 2007).

However, while looking at whether differences at the system level impacted student achievement by enabling better teaching and greater learning, McKinsey researchers found that the highest performing systems in the world, despite possessing large differences in culture, context, and construct, maintained a primary focus on instructional quality. These systems emphasized three things: (1) getting the best candidates into the teaching profession; (2) providing continuous, embedded in-service professional development; and (3) ensuring that the system responds to early signs of individual student failure.

Andreas Schleicher, head of the indicators and analysis division of the Organisation for Economic Co-operation and Development (OECD), which administers the PISA exam, drew an interesting conclusion after examining the McKinsey & Company findings. He observed that the high-performing systems shared a relentless focus on ensuring high instructional quality, while at the same time, reducing variability in the instruction every child receives (Schleicher, 2008). At McREL, we are particularly concerned about variability in system performance within schools, among schools and districts in the United States, and between U.S. educational systems and the rest of world.

Aiming for high quality, low variability

Most of the attention to achievement gaps in the United States has focused on the persistent performance differences among subgroups of students by race/ethnicity and socioeconomic status, particularly in our urban school districts. Increasing diversity in subgroup populations in suburban and rural districts has also contributed to achievement gaps being exposed that may not have been present before, or that were attributed previously to just a few students and overlooked. Additionally, as data collection and reporting systems have improved, districts and schools that were previously considered "high performing" began to find achievement gaps between subgroups of students that may have been previously masked.

Yet, in the report, *The Economic Impact of the Achievement Gap in America's Schools*, McKinsey & Company (2009) stress the importance of looking at two other gaps: (1) between similar students schooled in different systems or regions of the country; and (2) between the United States and other nations. In fact, "the most striking, poorly understood, and ultimately hopeful fact about the educational achievement gaps in the United States involves the huge differences in performance found between school systems, especially between systems serving similar students" (p. 12).

Possibly even more striking is the variability in instruction within schools. Hattie (2009) reports that, using multi-level modeling, researcher Spyros Konstantopoulos found a substantial proportion of the variation in student achievement lies within schools and not between schools. Grodsky and Gameron (as cited in Hattie, 2009) conclude that many of the influences that really make a difference to student learning in developed nations are within schools, from the influence of specific teachers, specific curriculum, and strategies teachers use to teach. Thus, one solution seems to be improving instructional quality while reducing the variability in the quality of that instruction within and among schools.

Defining system-level leadership and supports

Leadership plays a critical role in the performance of these "best in the world" systems. Indeed, the McKinsey & Company (2007) study notes that the research on school leadership suggests "school leadership is second only to classroom teaching as an influence on learning" (p. 29). Furthermore, they assert that school reforms rarely succeed without effective leadership, both at the level of the system, and at the level of individual schools. Researchers of another study noted that "there is not a single documented case of a school successfully turning around its pupil trajectory in the absence of talented leadership." (Leithwood et al., 2006, p. 5).

How do we define a "system of education"?

The majority of the OECD countries compared in the McKinsey & Company report have nationalized

education. In the United States, the school district is the legally defined entity for public education, but the devolution of "local control" varies from state to state. State departments of education and the U.S. Department of Education represent additional levels of scale, but they have incrementally less decision-making authority (although that has tightened substantially in the past three decades). Therefore, as we think about systemic solutions for system-level issues, we will focus primarily on the district and then look one level down to the school and up two levels to the state and national levels.

In the book, *District Leadership that Works: Striking the Right Balance*, Marzano and Waters (2009) present five district-level responsibilities from their meta-analysis that were found to be statistically correlated with increased student achievement. The research initially set out to answer the question of whether superintendent leadership in a district had an effect on student achievement. The findings from the study, however, indicate the importance of an expanded definition of district-level leadership, to include in addition to the superintendent, the collective central office staff, the board of education, and principals with their schools, operating as a school district versus a district of schools. The central theme across the responsibilities is the need to establish non-negotiable goals for achievement and instruction across the district as a whole.

Emulating High Reliability Organizations

Marzano and Waters (2009) went on to consider their findings about district leadership and defined autonomy from the perspective of High Reliability Organizations. Karl Weick has focused his research agenda in the last three decades to examining these organizations that "operate under high risk conditions and take a variety of steps in pursuit of error-free performance" (Weick, Sutcliffe, & Obstfeld, 1999). Weick's work with HROs began with studying flight deck operations on a nuclear aircraft carrier and has carried over to research including nuclear power plants, wildland firefighting, and aircraft flight operations.

In these industries, any mistake can have disastrous consequences—people die. To avoid disaster, these organizations put into place multilayered structures and processes to prevent errors and more importantly, to respond quickly before errors can cascade into catastrophic system failures. They also *mindfully* anticipate and manage the unexpected (Weick & Sutcliffe, 2001, 2007; Hoy & Sweetland, 2001; Hoy, 2003). Errors and mistakes are bound to occur, but the key is anticipating that they will occur and responding to them as soon as they appear. This constant monitoring for the early signs of failure and responding quickly is another way HROs demonstrate the characteristic of mindfulness.

At this point, a logical question to ask is, "What do these organizations have in common with K–12 public education systems and what can we possibly learn from them?" At McREL, we are translating the McKinsey & Company findings from the world's highest performing educational systems through a lens of high reliability. An operational definition of high reliability, applied to these systems, is this: *high levels of student performance, achieved as a result of high-quality instruction, delivered through superior execution of effective research-based practices, with low variability in the quality of instruction within and between schools.*

As we started to explore this idea of higher reliability educational systems, we ran across two other sets of educational researchers who were doing similar work. Sam Stringfield and David Reynolds began their theoretical exploration of HROs in 1991 (Stringfield, 1991) and, with the addition of Gene Schaffer, initiated a set of High-Reliability Systems (HRS) research studies in Great Britain in 1995. Stringfield and several of his colleagues have a long research history on the topics of teacher and school effectiveness and system improvement (Stringfield, 1991; Teddlie & Stringfield, 1993; Stringfield, Millsap, & Herman, 1998; Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002). Stringfield, Reynolds, and Schaffer approached their HRS project from an assumption that practices gleaned from these fields could be coupled with HRO concepts to establish a school improvement strategy.

Tom Bellamy and his colleagues also were examining the topic and wrote the article, "The Fail-Safe Schools Challenge: Leadership Possibilities for High Reliability Organizations" (Bellamy, Crawford, Huber-Marshall, & Coulter, 2005), in which they presented HRO as at least a metaphor, if not a model for education. They asserted:

> The stakes for failure have been raised so high … that high reliability has become an important aspect of school success. Schools are now challenged to prevent practically all failures and to close achievement (gaps) among student groups—in short, to ensure highly reliable learning for all students." (p. 384)

The hidden cost of underperformance

Although much attention remains focused on student failure (and rightly so), the ability of the United States to remain a global leader in innovation, science, technology, patents conferred, business, and social entrepreneurship will depend on the ability of its educational systems to

not only raise the floor, but also the ceiling. There are certainly moral and ethical aspects for closing the variety of achievement gaps, and severe social implications if we do not. Conversely, the economic benefits of dramatically raising the bar for the U.S. education system can be tremendous.

Economist Eric Hanushek's research has been not only on calibrating international assessments to one another for comparative performance, but also in constructing a sophisticated methodology for linking cognitive performance to economic growth, in terms of Gross Domestic Product (Hanushek, et al., 2008; Hanushek, Peterson, & Woessmann, 2010). Hanushek and his colleagues calculate if the United States had closed the gap between its educational achievement levels and those of other countries such as Korea and Finland, 2008 Gross Domestic Product could have been $1.3 trillion to $2.3 trillion higher (9%–16% of GDP). Furthermore, some see the persistence of these educational achievement gaps as impacting the U.S. economy with the equivalent of a permanent national recession (McKinsey & Company, 2009).

A case for urgency, a call for action

A move to high-performing, high-reliability, failure-free schools will, of course, require changes in cultures and systems. It will require that educators, policymakers, and the public examine basic assumptions about education in the United States and learn from the high performers, both here and abroad, to make America's schools among the best in the world.

On October 27–28, 2010, McREL, with support from the Kern Family Foundation, convened a small group of international thought leaders, forward-thinking superintendents, CEOs from educational organizations, and leaders from high-performance professions. Together, we explored the "new frontier" in improving the performance of U.S. schools and districts—a frontier that lies not in dreaming up new innovations or more "silver bullet" fixes for education, but rather, in flawless implementation of existing know-how to ensure all students benefit from top-quality instruction and learning environments.

At the Best in the World (BITW)[1] gathering, we particularly focused our attention on "the other achievement gap," the difference in performance between America's educational systems (even our highest performing districts) and those among the best in the world. The stated premises for this gathering were these:

- There is a gap in achievement between America's highest performing schools and school districts and the highest performing systems internationally.
- This gap in achievement may be a more serious threat to the future of the country than the gap between high-performing and low-performing U.S. schools and districts.
- This gap in achievement can only be closed by "raising the ceiling," or elevating the performance of America's highest performing schools and school districts.
- Elevating system performance, without excluding large numbers of students, requires a commitment to high performance with high reliability (raising the ceiling and the floor).
- Creating a constituency for, and urgency about, high-performing, high-reliability schools and districts is the biggest challenge we face in U.S. education.

The presentations from this group of experts (see sidebar on p. 5 for list of presenters) cemented the case that much can be learned from international comparisons, both from international comparative measures, such as PISA, and from benchmarking what works from those systems to school districts in the United States.

We understand the urgency, and we know that throwing large sums of money and a barrage of reform efforts at the problem hasn't resulted in significant, sustainable, or scalable change, but what is the best thing to do?

A theory of action—high performance with high reliability

While our focus for this gathering was on "the other achievement gap" between the United States and other systems of education worldwide, in light of Schleicher's conclusion, the variability in achievement found among, and within, state and district educational systems across the United States requires simultaneous attention. We believe that lessons learned from High Reliability Organizations may provide us with a foundation for school improvement and with a set of principles and strategies to directly apply to our educational systems. We asked Stringfield and his colleagues along with Tom Bellamy, to help us develop a theory of action.

Bellamy and Stringfield also prepared commissioned papers to accompany their presentations at the Best in World Exploratory Gathering. Those two papers constitute the next two chapters of this monograph. We invite you to read on.

[1] For video clips of the presentations, visit the Network for Innovative Education website at https://sites.google.com/site/networkforinnovativeeducation/Home.

Best in the World presenters

- Sir Michael Barber, expert partner in McKinsey & Company's Global Public Sector Practice and head of Global Education Practice, and co-author of *How the World's Best-Performing School Systems Come Out On Top*.
- Martin West, deputy director of the Program on Education Policy and Governance, Harvard University
- Tom Bellamy, professor and director, Goodlad Institute for Educational Renewal, University of Washington Bothell
- Sam Stringfield, professor and distinguished university scholar, University of Louisville; with Eugene Schaffer, professor and chair, Dept. of Education, University of Maryland Baltimore County; and David Reynolds, professor, University College Cardiff, Wales

References

Bellamy, T., Crawford, L., Huber-Marshall, L., & Coulter, G. (2005). The fail-safe schools challenge: Leadership possibilities from high reliability organizations. *Educational Administration Quarterly, 4*(3), 383–412.

Hanushek, E. A., Jamison, D. T., Jamison, E. A., & Woessmann, L. (2008, Spring). Education and economic growth. *Education Next*, 62–70.

Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2010, November). *U.S. math performance in global perspective: How well does each state do at producing high achieving students?* Retrieved from http://www.hks.harvard.edu/pepg/PDF/Papers/PEPG10-19_HanushekPetersonWoessmann.pdf

Hattie, J. A. C. (2009). *Visible learning: A synthesis of meta-analyses relating to achievement*. New York, NY: Routledge.

Hoy, W. K. (2003). An analysis of enabling and mindful school structures: Some theoretical research and practical considerations. *Educational Administration Quarterly, 41*(1), 87–108.

Hoy, W. K., & Sweetland, S. R. (2001). Designing better schools: The meaning and measure of enabling school structures. *Educational Administration Quarterly, 37*(3), 296–321.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). *Seven strong claims about successful school leadership*. London: Department for Education and Skills.

Marzano, R., & Waters, T. (2009). *District leadership that works*. Bloomington, IN: Solution Tree.

McKinsey & Company. (2007). *How the world's best performing school systems come out on top*. New York, NY: Author.

McKinsey & Company, Social Sector Office. (2009, April). *The economic impact of the achievement gap in America's schools*. Retrieved from http://www.mckinsey.com/app_media/images/page_images/offices/socialsector/pdf/achievement_gap_report.pdf

OECD Programme for International Student Assessment (PISA). (2010). *The high cost of low educational performance*. Paris, France: Author.

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, E. (2002). *World class schools: International perspectives on school effectiveness*. New York, NY: Routledge/Falmer.

Schleicher, A. (2008, June). *Seeing the U.S. school systems through the prism of international benchmarks*. Paper presented at the Annual Policy and Practice Institute for School Leadership, Dover, DE.

Stringfield, S. (1991, January). *A model of elementary schooling effects: Evidence of support*. Paper presented at the meeting of the International Congress for School Effectiveness and Improvement, Cardiff, Wales.

Stringfield, S., Millsap, M., & Herman, R. (1998). Using "promising programs" to improve educational processes and student outcomes. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *International Handbook of Educational Change* (pp. 1314–1338). Dordrecht, The Netherlands: Kluwer.

Teddlie, C., & Stringfield, S. (1993). *Schools make a difference*. New York, NY: Teachers College Press.

Walberg, H. J. (2003, Spring). Accountability unplugged. *Education Next,* 76–79.

Weick, K. E., & Sutcliffe, K. M. (2001). M*anaging the unexpected: Assuring high performance in an age of complexity*. San Francisco, CA: Jossey-Bass.

Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty*. San Francisco, CA: Jossey-Bass.

Weick, K. E., Sutcliffe, K., & Obstfeld, D. (1999). Organizing for high reliability: Process of collective mindfulness. *Research in Organizational Behavior*, 21, 81–123.

Chapter Two

Toward Highly Reliable, High-Quality Public Schooling

By Sam Stringfield, University of Louisville; David Reynolds, University of Southampton; and Eugene Schaffer, University of Maryland, Baltimore County

Introduction

Large numbers of U.S. schools and their students can achieve at levels fully competitive with, and perhaps redefining, "the best in the world." Furthermore, empirical evidence suggests that dramatic improvement in American educational reliability—educating virtually all children to fundamentally higher levels of measured achievement—is possible.

The challenge of providing "the best in the world" education involves knowing "what works extremely well" and providing it with remarkable reliability. These two components are multiplicative; that is to say, the quality of education received by students in any given classroom, school, local education authority (LEA), state, or nation is the product of the extent to which the providers use state-of-the-art, proven methods and processes multiplied by the reliability of delivery. This idea, stated as a straightforward equation, becomes this:

> Effectiveness of schooling = Effectiveness of the "Technology" x Reliability of Delivery

Much more—and typically more rigorous—research exists on the subject of "what can be effective" than on "how to reliably deliver it." To partially re-balance this situation, our focus in this chapter is on methods for improving the reliability of educational reform efforts.

Trend data and implications

High Reliable Organizations evolve only when the professionals working in an area and the larger public come to believe that the historic levels of the organization's reliability are likely to lead to disaster. The data we examined and present here indicate that the United States is moving precisely toward that condition. We observe, too, that in such complex systems, teachers and other educators see a "loosely coupled system" (Weick, 1976) and respond, for purposes of their long-term professional survival, as "street level bureaucrats" (Lipsky, 1980). Stated in more theoretical terms, when fundamentally interesting, potentially valid reforms have been attempted in parts of the complex education system, they were predicted to fail—and often did—not because the ideas were invalid, but because they were overwhelmed by the larger logic of the system. A point to which we will return repeatedly is that in educational reform, as in research, reliability sets the upper boundary of measured validity.

International trend data

The Trends in International Mathematics and Science Study (TIMMS)[1] provides an unparalleled set of studies spanning more than a decade and providing comparisons and contrasts of student achievement among nations' systems of schooling. While lacking the breadth of TIMSS, Reynolds, Creemers, Stringfield, Teddlie, and Schaffer (2002) produced a mixed-methods study contrasting higher and lower achieving schools within and across nine countries' educational systems. Both of these studies examine the comparable value of schooling on student test performance. The first looks at cross-national data while the second looks at within-country variance and cross-country variance following cohorts for two years.

The National Center for Educational Statistics (NCES) regularly provides reviews of various international comparisons of student achievement (TIMSS, the Progress in International Reading Literacy Study [PIRLS], and the Program for International Student Assessment [PISA]). NCES reported that in 4th- and 8th-grade reading, U.S. students ranked 10th of 45 nations/provinces studied. Disturbingly, the NCES review found that over time, a growing number of countries' students were exceeding the average of U.S. students in reading abilities. In 4th- and 8th-grade mathematics, the TIMSS data indicate that U.S. children are making progress over time relative to students in other countries, and

[1] For an overview and range of reports, see http://nces.ed.gov/timss.

currently rank in the top 8 of over 40 countries on mathematics measures.[2] In the 2006 PISA science study, U.S. 15-year-olds scored in the bottom third of OECD (economically developed) participating nations. To the extent that our goal is to be "the best in the world," we have some distance to travel.

U.S. longitudinal achievement trends

Complaints about "the current state of schooling" in the United States and the need for "dramatic improvements" have been staples of the American political scene for over 200 years. Consider that at the end of the 19th century, Harvard's president complained that the American students at college entry simply were not up to European standards. Eliot (1898) laid the blame squarely on American schools, whose "main characteristic of instruction is dullness, a complete lack of human interest and a consequent lack in the child of the sense of increasing power" (p. 184, as cited in Nunnery, 1998). Nearly a century later, John Goodlad (1984), summarizing a large study he and a team had completed, observed, "Only rarely did we find evidence to suggest instruction likely to go much beyond mere possession of information…Boredom is a disease of epidemic proportions" (pp. 236–242).

We posit the sameness of the critiques is not the result of lack of change efforts and offer a few facts regarding educational outcomes over the last century, the relative stability over the last 30 years, and the costs of that sameness in an ever-changing and educationally improving world:

1. The overall percentages of Americans per birth cohort who graduated from high school stood at approximately 20 percent at the dawn of the 20th century, and rose to 76 percent in 1970. The high school graduation rate then gradually declined to 68 percent in 1998, and subsequently has risen steadily to its current 75–77 percent today. The good news here is that in the first decade of the 21st century, the rise in the percent of students graduating from high school is as steep as at any time in our nation's recorded history (Heckman & LaFontaine, 2010). The bad news is that a quarter of our young people are leaving education almost completely unprepared to compete in a global 21st century information economy.
2. Of our non-high school graduates, the major change has been in the percentage taking and passing the General Educational Development tests (GEDs), which is considered to be a high school equivalency certificate. The percentage of high school dropouts taking the GED tests has increased rapidly over the past 15 years and helps explain the substantial gaps between U.S. Census data on percentages of young adults who are "high school graduates" (including, for Census purposes, GED holders) and data on actual graduation rates. Unfortunately, today the GED provides very modest economic value over high school dropouts not completing a GED. By contrast, high school graduation has approximately a 50 percent "value added" over not graduating.
3. The percentages of Americans with four-year college degrees or higher has risen steadily throughout the last century, from approximately 5 percent of the cohort born in 1900 to approximately 31 percent today. Interestingly, the nation's gains over the last 30 years in percentages of young people graduating from college have been the result of increasing percentages of high school graduates—especially females—attending and completing college.
4. For nearly four decades, the U.S. Department of Education has conducted extensive, nationally representative studies of student achievement in the areas of reading and mathematics. These data comprise the National Assessment of Educational Progress (NAEP) long-term trend data.[3] Figure 1 provides data, presented in mean scale scores, on NAEP student reading scores at ages 9, 13, and 17 from 1971 through 2008. The relatively good news in Figure 1 is that age 9 reading mean scores are at the highest level yet measured on NAEP. The less encouraging news is that the age 13 scores are at the same level as in 1992, and are not dramatically above the levels of 1971. The least encouraging news is that the mean score for 17-year-olds is not statistically different from the scores in 1971 and is actually significantly below the scores from the mid-1980s through early 1990s.
5. Figure 2 provides NAEP mean mathematics scores for 9-, 13-, and 17-year-olds from 1973 through 2008. Again, the most encouraging news is from the 9-year-olds. Those scores have risen significantly and at an impressive rate of progress over the last 35 years, with the steepest rise coming in the last decade. Statistically significant and only moderately

[2] See http://nces.ed.gov/programs/coe/2009/analysis.

[3] See http://nces.ed.gov/nationsreportcard/ltt.

Figure 1

Trend in NAEP Reading average scores for 9-, 13-, and 17-year-old students, 1971–2008

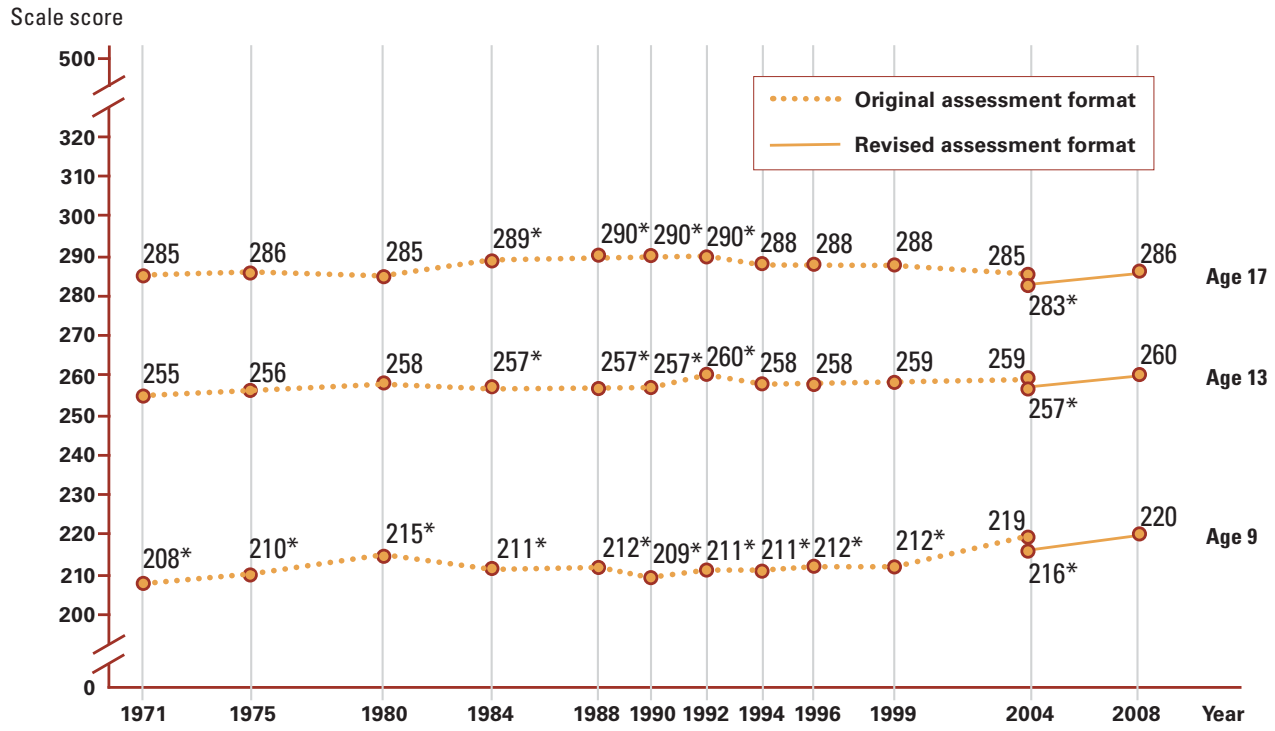
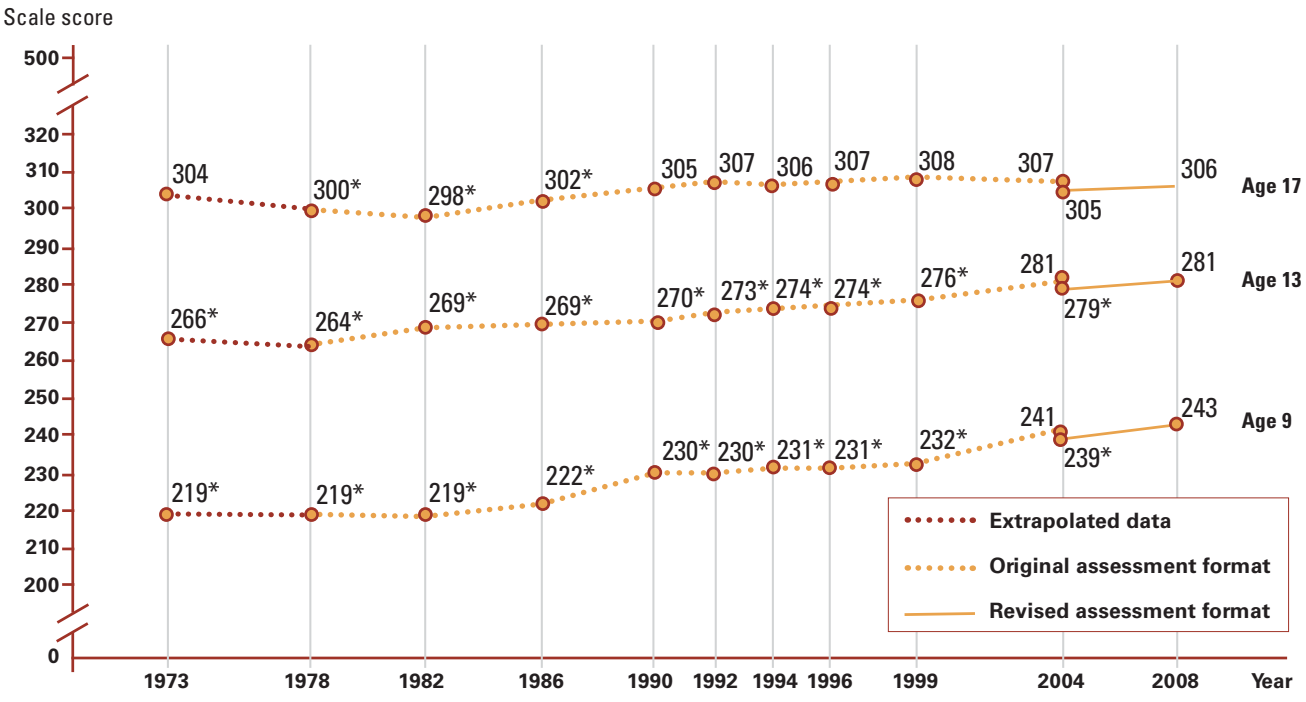


Figure 2

Trend in NAEP Mathematics average scores for 9-, 13-, and 17-year-old students, 1973–2008



less impressive progress has been made by our schools in advancing the mathematics achievements of 13-year-olds. However, as with reading scores, 17-year-olds' mean math scores in 2008 were not statistically different from scores in the early 1970s and do not reflect a significant rise in the last decade.

In summary, our level of high school graduation has risen for most of the last century, fell modestly for two decades, and is again at its highest levels in history, at approximately 75 percent. Including holders of GED certificates, over 85 percent of young adult Americans are counted as being "high school graduates." Rates of college attendance and graduation have risen almost continuously through the last century and today stand at 31 percent of each new birth cohort. The best available evidence of long-term reading and mathematics achievement among American school-aged students shows clear 30+ year gains in the elementary grades, but no appreciable gains by the upper grades of high school, as students prepare for college and careers, or both. So, what does this mean, exactly?

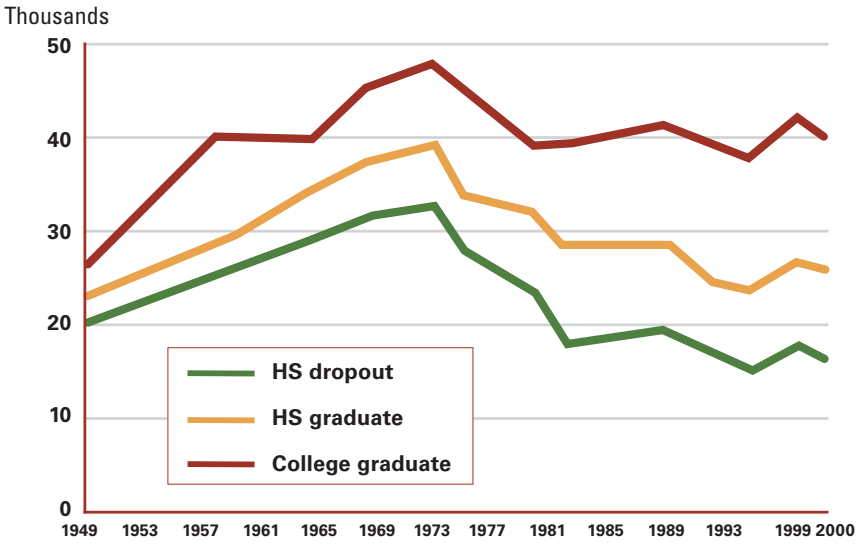
Economic and other impacts of education in the United States

Figure 3 presents data from 1949–2000 on the median income of young adult (age 25–34) American males. We focus on young adults because the effects of education are first noticeable in the age range that could be expected to have completed their formal education. We focus on males because in the late 1940s only 25 percent of U.S. females were working outside the home, and that number had risen to 80 percent by the year 2000, hence making whole-cohort comparisons among females over time problematic. One of the most striking sets of facts revealed in Figure 3 is the change in the economic benefit of obtaining additional education. In 1949, the average young male college graduate made 30 percent more in annual income than the average high school dropout. (In other words, for every dollar the average young male high school dropout earned, a college graduate earned $1.30.) By the year 2000, that differential had expanded to over 150 percent. (For every dollar earned by a young male dropout, the college graduate made over $2.50.) The economic advantage of succeeding in schooling had increased by over five-fold. In constant (inflation-adjusted) dollars, a 21st century high school dropout not only makes less than his grandfather, the high school dropout, made after World War II, he makes less than half as much as his father, the high school dropout, made in the early 1970s.

For a current picture, we need not limit the discussion to males or young people. Figure 4 presents a point-in-time data set of all Americans' (male and female, all ages) income by education in 2008. According to the U.S. Bureau of Labor Statistics, across all persons in the would-be-working range of our population, the probability of not being able to find gainful employment was over three times higher for high school dropouts than for college graduates, regardless of age or gender. Regarding income, for every dollar the average high school dropout earned, the average high school graduate made $1.38, the average college graduate made $2.30, and the average professional (Ph.D., M.D., etc.) made over $3.50. Discrepancies of this magnitude were almost unimaginable in the late 1940s and 1950s, but they are today's realities. Further expanding the differences, the typical college graduate marries another college graduate, with the practical implication being that the differences in family income by education often are doubled. Finally, in 2010, for the first time in U.S. history, more women are working outside the home than

Figure 3

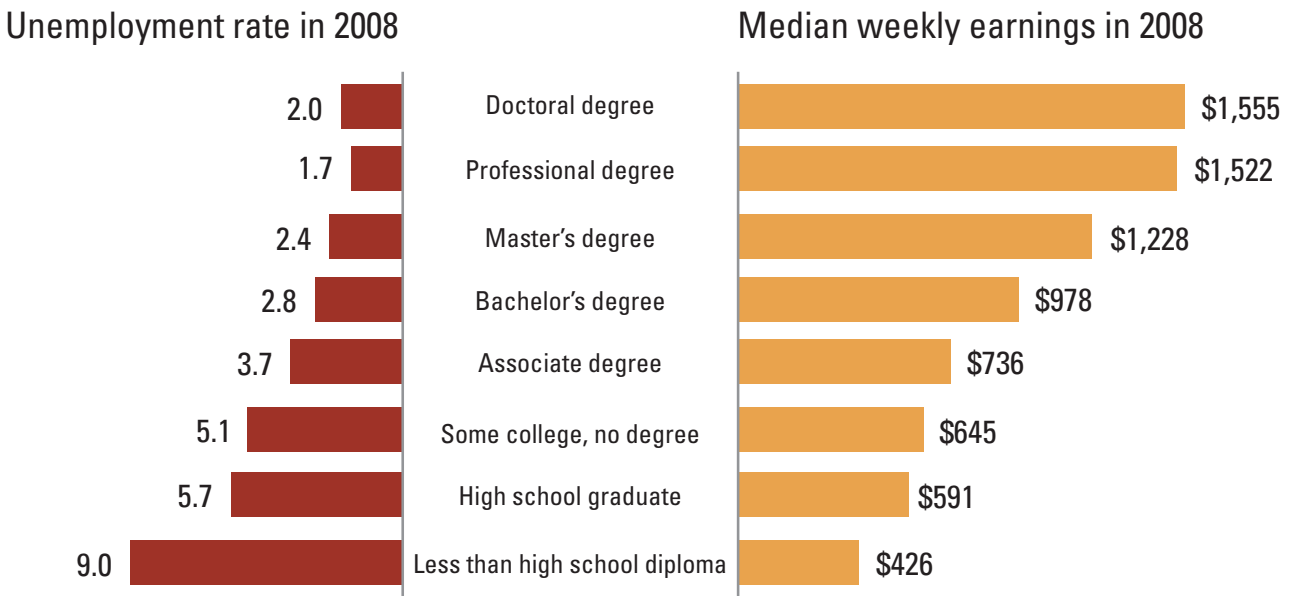
Median income of male Americans aged 25–34, by educational level, 1949–2000



Source: Educational attainment in the United States, U.S. Census Bureau, March 2000. 2000 dollars (CPI-U adjusted)

Figure 4

Education pays



Source: Bureau of Labor Statistics, Current Population Survey

Note: The BLS unemployment statistic averaged 6.2% in 2008. In September of 2010 it stood at 9.7%.

men, and more women are working in managerial positions.

We do not believe that income and employment are the most important variables in life, only that the longest span of reasonably reliable data are available on them. Regarding other variables, a range of studies have found that persons with higher levels of educational attainment are more likely to engage in a wide series of additional, pro-social activities. Whether the topic is voting, taking leadership positions in organizations ranging from churches to political parties, starting new businesses or staying off welfare and out of prison, increases in education are consistently correlated with success in life. The conclusion of studies of income and of a range of other desirable adult outcomes must be that success in school now matters more than at any other time in our history. Today, educational failure is catastrophic for the individual, his or her future family, their community, and our society. So, why is it that seemingly logically valuable reform efforts have tended to not produce positive results?

Complex, loosely coupled systems, and the inevitable roles of "street-level bureaucrats"

Education in the United States is a very complex, loosely coupled system. This became abundantly clear to Stringfield when he was appointed to the New Board of School Commissioners of the Baltimore City Public Schools System (BCPSS) in 1998. He soon found the challenges bewildering (see Stringfield & Yakimowski, 2005; Stringfield, 2008). A particular point of importance—and frustration—in his experience was the search for credible levers for affecting change from the seemingly powerful position of the school board. This led to a series of attempts to model the relationships among various levels of the modern American education system (Datnow, Lasky, Stringfield, & Teddlie, 2006; Land & Stringfield, 2005).

No shortage of examples of complex educational systems

Figure 5 (see p. 12) presents Stringfield's (2005) representation of the complex set of relationships among levels of educational governance in the United States. To take one practical example of these relationships, consider the passage of the federal No Child Left Behind (NCLB, 2001) legislation. Congress passed the legislation, but the U.S. Department of Education took over a year to develop regulations and even longer to develop non-regulatory guidance. (With federal educational legislation, states more often look to the non-regulatory guidance to determine how to meet legal requirements.) States then developed new testing schemes and established new regulatory requirements of their own, which they passed on to LEAs.

Meanwhile, colleges of education across the country were changing requirements in various courses and developing new programs to assist schools and districts in meeting the new requirements. More aggressively, a broad range of for-profit corporations (such as text and software publishers and consulting firms) and not-for-profit entities (such as the regional laboratories and various foundations) began developing products, workshops, and other materials to assist schools and LEAs in addressing the changes required in NCLB. LEAs received this range of information and federal funds, and made diverse new requirements on schools, which in turn made new demands on teachers. The theory of action here is that federal laws, coupled with funding that totals less than six percent of the average districts' budgets, will produce substantial change at each subsequent level, eventually resulting in increased student achievement. At the level of Figure 5, this presumes a tidy, tightly coupled system for educational improvement.

However, even the slightest examination of what happens within any one of the components of the system, let along across the full set, tells a different story. Figure 6 follows the formal, official model of how changes happen within just one box of Figure 5. The one box is the LEA, but any other would make the same point. Within Figure 6 (see p. 13), a school board—typically elected, but in some large systems appointed—considers the policy implications of new laws and other inputs, sets policy, and directs the superintendent (the board's one employee) to implement policy. The superintendent works through the various offices under his or her control (deputy superintendents for each school level, professional development, accounting, and testing, among others), and that group works with schools to implement the mandated-from-above changes. Goals are set and communicated, special programs are created or re-directed, and standardized tests and other measures provide feedback loops. Clear lines of authority exist within LEAs, and in theory, these are relatively tightly coupled.

Figure 7 (see p. 13) presents something closer to Stringfield's school board experience at the LEA level. The board did all of the things that the presumed theory of action dictated: considered options, established goals, directed the flow of money (while checking to be sure that the budget balanced), agreed on measures, and provided clear direction to the superintendent. The superintendent met with his district leadership team, discussed tactical options, passed down practical strategies, and so on. But almost immediately, a complex series of loops appeared.

Examples were everywhere. A board member's long-term neighbor and trusted friend (a teacher or a principal) came to the board member's home to complain bitterly. A third-level functionary somewhere in the central office discovered an inconsistency between new policies and old ones that had guided practical actions for 10 years. The reallocation of Title I funds was perceived by an elected official's spouse to disadvantage their 5th-grade child. A politically well-connected principal who had, over years, garnered great community support for her school either ignored the new directives with impunity or, if pushed, organized formal opposition and demanded change to what had been. End-of-year test scores came in, and they did not demonstrate dramatic short-term effects of the by-now-unpopular changes. What quickly become apparent were the second-, third-, and fourth layers of formal and informal communications and powerful counter-veiling change forces.[4] Almost immediately, the theoretically somewhat tightly coupled systems proved to be loosely coupled, calling for a different theory of action and related strategies for change.

"Loose coupling" is a term Karl Weick (1976) used to describe the working of schools, and we will return to it shortly. First, we briefly describe one aspect of educators' worlds, and their "street-level" policy implementation.

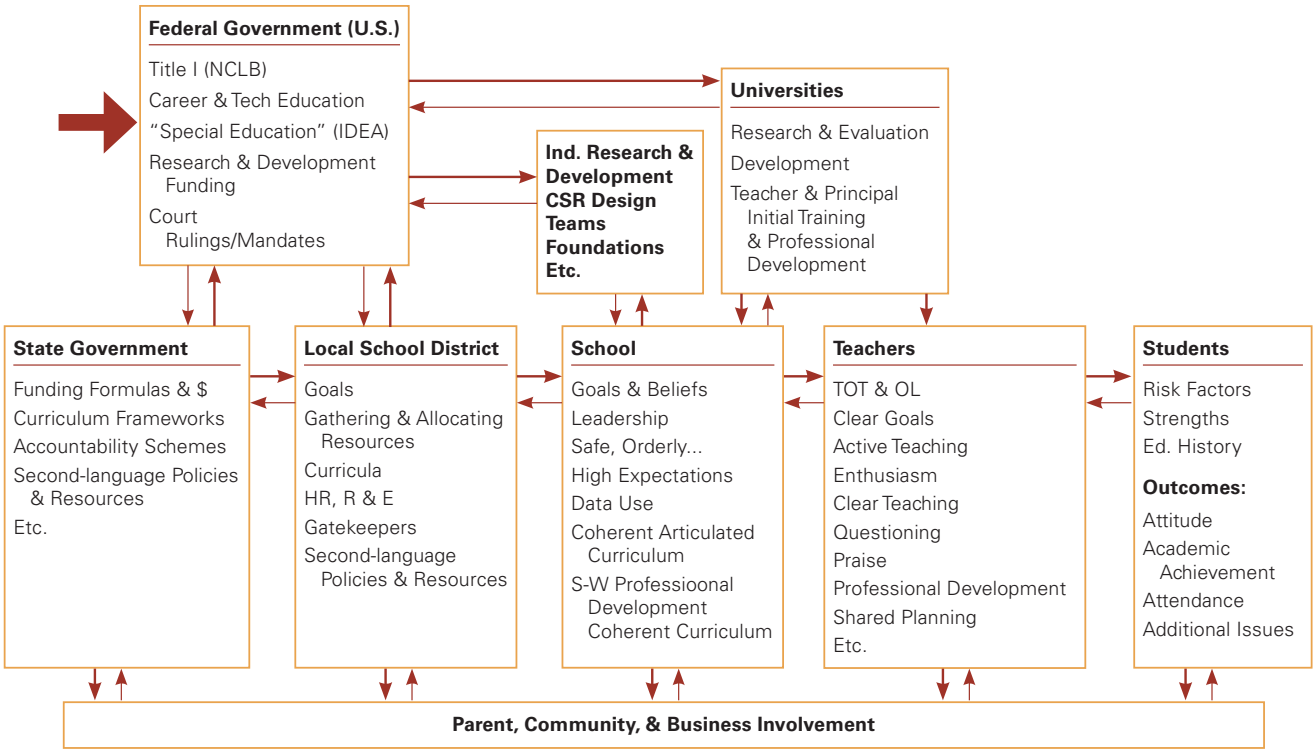
Street-level bureaucrats

Teachers, principals, and local educational administrators are necessarily practical people. Faced with simultaneous requirements to continue work in their classrooms (schools, district offices, etc.) and to implement changes, the full range of implications of which have almost necessarily not been thought through by those making the new demands, practical educators necessarily behave as "street-level bureaucrats" (Lipsky, 1980). Lipsky described autonomous workers, such as police officers and social workers, as working in arenas enmeshed in vague and often conflicting goals, accountability requirements, large demands for services—often more than an individual can provide, and often to involuntary clients—and the additional requirement of performing with limited and typically inadequate resources. These employees must find ways to manage under requirements that, if taken literally, would be impossible for any one human being

[4] As this chapter is being written, opponents of Louisville's (Ky.) long-standing student busing system are simultaneously in court arguing for a 100 percent return to neighborhood schools and mobilizing support to vote out board members who have supported the policy that once had been mandated by federal courts. Change forces are more complex than most of us realize until we try to change something.

Figure 5

A static representation of the relationships of educational organization levels and their potential influences on students



to implement. So, these street-level bureaucrats negotiate the space in ways they individually deem best. Policemen decide which of the thousands of laws to enforce as they walk their beat; social workers decide which clients to turn in for minor infractions, and when to look the other way.

Faced with requirements to maintain at least a minimum level of classroom discipline; collect lunch money; prepare lessons; write, administer, and score tests; and literally hundreds of other, frequently changing tasks, educators, like all other street-level bureaucrats, make choices as to which policies, old and new, to implement this hour, day, and year. Hardly surprisingly, over time they develop sophisticated personal and group systems for filtering and interpreting new requirements that may or may not be achievable when added to current tasks. In short, educators at several levels work in very complex, often conflicting environments. They must interpret each new signal and decide whether they can respond, and if so, determine how best to do so. It is at this "street level" that educational change does or does not become a reality.

Loose coupling

Karl Weick (1976) observed that educational organizations are "loosely coupled systems" (p. 1) and noted several advantages. On the upside, loose coupling allows some portions of an organization to persist. Loose coupling prevents each part of an organization from having to respond to every single new signal in a system (i.e., no one has to do the impossible continuously).

In addition, loose coupling allows for localized adaptation. Persons and groups face different challenges and are often best served by addressing their situations differently. These not-centrally-planned mutations sometimes eventually prove valuable to the larger whole, and certainly they can facilitate local functioning. (However, loose coupling makes it much harder for an organization to change as a whole unit.) If there is a breakdown in one portion of a loosely coupled system, the breakdown need not affect other parts of the system. Given that people enjoy holding a sense of self-control and self-efficacy, it is noteworthy that loose coupling generates autonomy. This enhances individuals' sense of self-determination and may raise morale.

Regarding cost, Weick hypothesized that loosely coupled systems "should be relatively inexpensive to run, because it takes time and money

Figure 6

The local school district

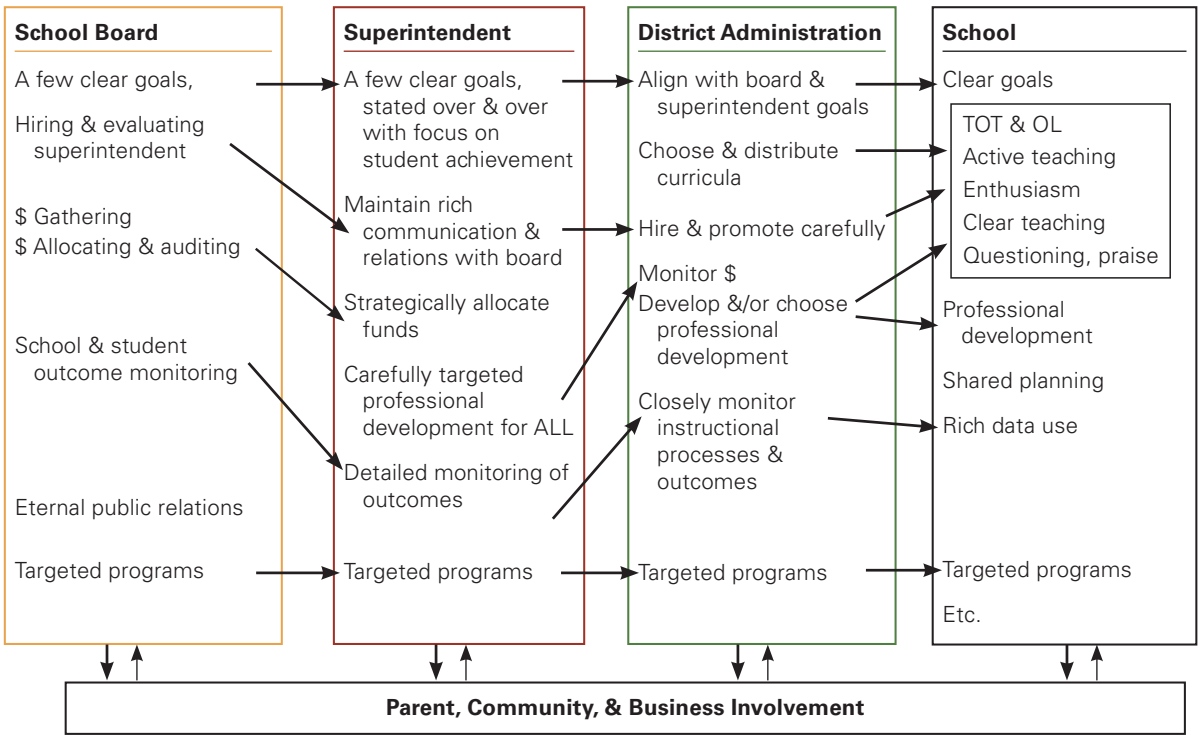
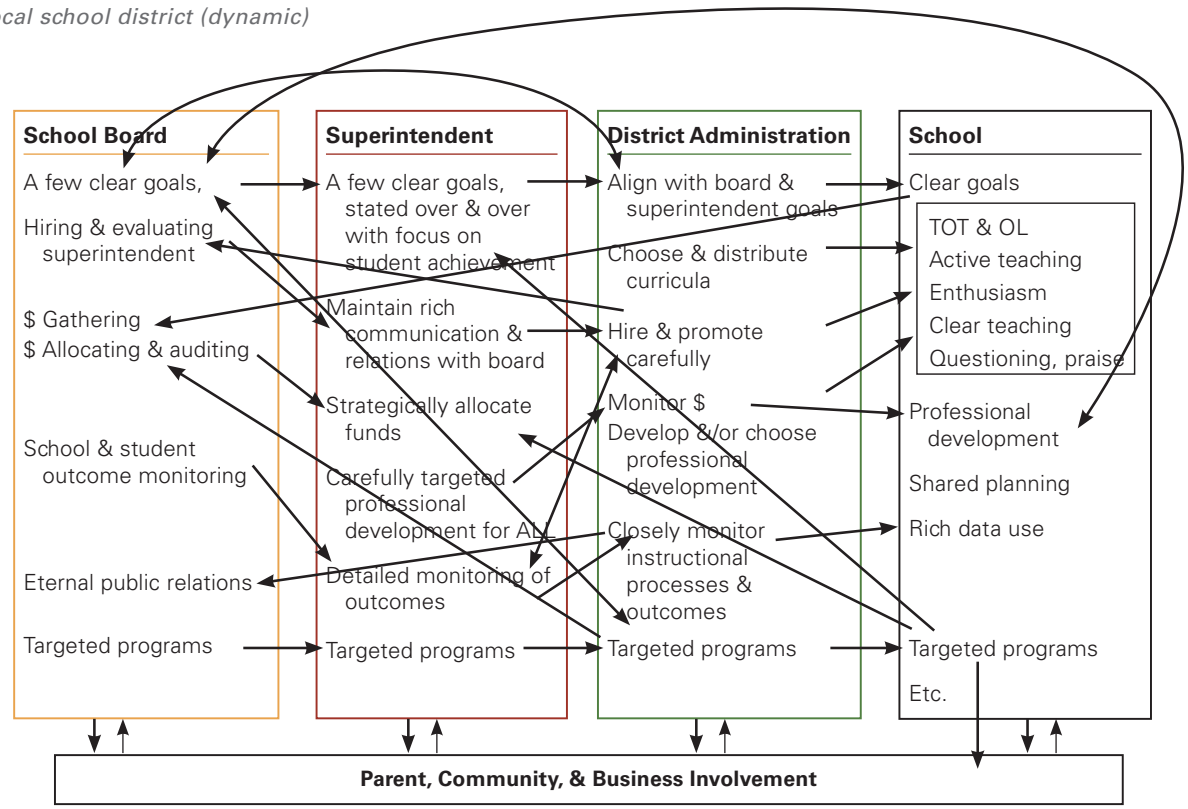


Figure 7

The local school district (dynamic)



to coordinate people" (p. 8). The tradeoff for being inexpensive is that loose coupling produces non-rational systems of fund allocation.

Finally, Weick observed that no organization could be tightly coupled in all areas. Tight coupling in some areas requires loose coupling in others. The ratio of tight and loose coupling varies among organizations.

Weick was neither exclusively pro- nor anti-loose coupling. His goal was to describe "what is." An oft-expressed, street-level way of showing Weick's central points is the image of "the egg carton school" in which each teacher may close the classroom door and do as she or he pleases in relative isolation and independence. The same egg carton model has historically described the relationships among schools and between schools and LEAs, LEAs and states, and the 50 state departments of education and the U.S. Department of Education.[5]

We believe that the combination of teachers, schools, and school systems working in loosely coupled relationships, and of educators necessarily making street-level decisions on an ongoing basis describes much of modern American educational practice. Among educators, there often is a firmly held attitude that research has little to nothing to offer practical educators[6] and that most school systems have a new "focus" for each new school year, therefore ensuring the current focus will be gone next year, if not sooner. Loose coupling, when combined with these two commonly held beliefs, explains much of the failure of various educational reform efforts to gain traction within and across schools.

Complex interventions inserted into inadequately understood, deeply complex, loosely coupled systems are unlikely to have measurable effects, even if they intermittently reach the street-level bureaucrats charged with implementing them. Under-funded, under-supported, inadequately trained street-level bureaucrats have few choices other than to keep moving to the next mandate. Given such realities, and remembering that reliability sets the upper boundary of measured validity, nearly all of us would predict that almost all reform efforts would fail.

Serendipity favors the prepared minds: The evolution of High Reliability Schools (HRS)

In the summer of 1989, Stringfield was trying to solve a data riddle. He and Charles Teddlie had been analyzing a mountain of quantitative data gathered in the Louisiana School Effectiveness Study (LSES).[7] LSES phases three and four had included a "double blind." The 16 schools in those phases were carefully chosen matched pairs. In each demographically matched pair, one had a multi-year history of unusually high academic performance, and the other school had a stable history of underperformance. Neither the schools nor the observers were to know which school of each pair was the "positive" or the "negative" outlier.

We eventually learned that one aspect of the design had failed: armed with no prior knowledge, 100 percent of the observers had intuited the status of 100 percent of the schools where they had observed. Stringfield read and re-read the observers' detailed case studies, trying to understand how the observers—most of whom were not professional educators and some of whom had not been in an elementary school since they were students—had done it. Every negative outlier school had at least one exceptional educator, and many had multiple attractive characteristics. None of the positive outliers was implementing "the latest" reform, and in fact, several looked at first blush to be as plain vanilla as schools could be.

As he ponderously distilled eight dimensions of qualitatively observed differences (eventually published as Stringfield & Teddlie, 1991), he chanced to read the then-current edition of *Smithsonian* magazine, which included a popularized article on High Reliability Organizations (Pfeiffer, 1989). HRO contained a vocabulary for explaining what was being described in the LSES case studies: in the negative outliers, a wide range of behaviors and results were tolerated, while the positive outlier schools had in common a clear focus on students' academic achievement and intolerance for observably ineffective educator behaviors. Contrasted with the negative outliers, the positive outlier schools were much more reliable education providers.

Here are the 12 HRO principles, briefly stated and seen through the eyes of an educational effectiveness researcher:

1. Organizational reliability evolves under a particular circumstance. HROs evolve when both the larger society and the professionals involved in the

[5] In the mid-1980s, Stringfield managed Northwest Lab's Chapter 1 Technical Assistance Center in Denver, working with Chapter 1 (Now Title I) programs across several western states. The work required regularly moving between providing professional development to teachers and para-professionals within and across states, working with LEAs and SEAs, and semi-annual meetings with federal officials in Washington, D.C. As would be predicted by research on loosely coupled street- (and federal-) level bureaucrats, Stringfield quickly learned "the Chapter 1 law," as implemented, varied greatly among states, among LEAs within states, and among schools within LEAs.

[6] We believe that universities are at least partially responsible for this issue. We in universities often require few-to-no courses in learning how to differentiate between credible research and opinion, and when we do offer the courses, the practical applicability of what is offered is often limited.

[7] The quantitative and some of the qualitative results were published in a series of articles and in Teddlie and Stringfield (1993).

working of the organization come to believe that failure of the organization to achieve its key goals would be disastrous. (As noted previously, we believe that this condition is rapidly being met today.)

2. HROs require a clear and finite set of goals, shared at all organizational levels.

3. An ongoing alertness to surprises or lapses exists, and small failures in key systems are monitored closely because they can cascade into major problems. In order to sustain multi-level awareness, HROs build powerful databases. These databases possess "Four R's": relevance to core goals; rich triangulation on key dimensions; real-time availability to all organizational levels; and regular cross-checking by multiple, concerned groups.

4. The extension of formal, logical decision making analysis as far as extant knowledge allows. Regularly repeated tasks that are effective become Standard Operating Procedures (SOPs).

5. HROs actively sustain initiatives that encourage all concerned to identify flaws in SOPs and honor the flaw finders.

Because high reliability is a social construction and requires high levels of individual professional decision making, HROs perpetually engage in the following three activities:

6. Active, extensive recruiting of new staff at all levels.

7. Constant, targeted training and retraining.

8. Rigorous performance evaluation.

Four additional characteristics follow:

9. Key equipment is kept in high working order.

10. Because time is the perpetual enemy of reliability, HROs are hierarchically structured. However, during times of peak activity, whether anticipated or not, HROs display a second layer of behavior that emphasizes collegial decision making, regardless of the formal position of the decision maker.

11. Clear, regularly demonstrated valuing of the organization by its supervising and surrounding organizations. All levels work to maintain active, respectful communication geared to the key goals of the HRO.

12. Short-term efficiency takes a back seat to very high reliability.

Two additional points relate to the HRO characteristics. The first is that while these characteristics must necessarily be described separately, *their effects are presumed to be multiplicative, not merely additive*. The total absence of any one can nullify great efforts to obtain others. Standard Operating Procedures can become mindlessly rigid in the absence of ongoing honoring of flaw-finders and process/program improvers. Aggressive recruiting in the absence of supportive, long-term professional development is futile. The first 11 characteristics cannot be sustained if an organization continues a history of such poor accounting and economic prediction that it must periodically make drastic cuts in personnel, equipment, etc.

A second note concerns the description of the characteristics. It would be easy to regard each of the above HRO characteristics as existing in a stable state. In fact, all are *dynamic and regularly evolving*. As technologies advance, systems have the opportunity to create much richer databases. Last year's teacher recruiting effort, however successful, becomes the baseline for measuring this year's effort, and so on. In human organizations, reliability is a socially constructed, evolving phenomenon.

The High Reliability Schools project is born

In 1991, Stringfield wrote a "think piece" exploring the potential for HRO principles to be used in school reform and presented it at the International Congress for School Effectiveness and School Improvement (ICSEI). David Reynolds, the founding co-editor of *School Effectiveness and School Improvement*, happened to be in the audience. A year later, Reynolds was presenting an after-school lecture to a group of British educators and briefly discussed the fact that a "mad American" had the idea that schools could be operated with the same reliability as air traffic control towers. To Reynolds' surprise, a group of educators came to him immediately after his presentation and said, "Let's do it." Reynolds asked, "Do what?" The local educators stated a desire to try to operate their schools with the remarkable reliability of air traffic controllers, and the High Reliability Schools project began.

HRS's immediate challenge was to take the abstract ideas of HROs and convert them into concrete, usable professional development segments for teachers and "heads" (in the United States, principals). As the boundaries of Stringfield's and Reynolds' skills in this area became manifest, the two contacted a colleague, Gene Schaffer, who was exceptionally skilled in translating abstract ideas regarding "effectiveness" and "school improvement" research into concrete educational professional development segments. This team has now worked together for nearly 15 years.

Three overarching sets of ideas defined the HRS project. The first was that the broadly defined fields of teacher, school, and system

effectiveness had evolved far enough to provide some level of guidance to practical educators. The second concerned our conception of the components of HROs.[8] Third, our assumption was that the specifics were not so refined that they could be implemented lock-step in every school and classroom, but that the general principles would be able to guide local educators who were willing to work with college professors to "co-construct" (Datnow & Stringfield, 2000) a reform. We were explicit with all the local educators that we would be entering as equal partners in an exploration, or not at all. We professors would bring relevant research knowledge to the table, and the diverse local educators would work with us to make practical applications of the research.

The research knowledge bases we attempted to bring to the HRS schools were these:

1. **Teacher effects.** For example, Good and Brophy's (1987, 2007) *Looking in Classrooms*, including training in a series of low- and high-inference teacher-effects-related observation instruments.

2. **School effects.** For example, including broad, general principles (e.g., "school climate") and specific, alterable variables (e.g., "effective use of class time").

3. **Extensive data gathering and use.** This was a relatively under-developed field at the time, but we encouraged schools to adopt a testing scheme that was new, and promised to provide early indications of students' ability to perform on national tests at age 16.

4. **Popularized findings from studies of improving businesses.** Such things as adopting "big, hairy, audacious goals" (BHAGS) (Collins & Porras, 1996) instead of modest, "reasonable" goals.

We believed that the sets of findings from these fields, although valid, were being implemented around the United States and in various locations around the globe with indifferent fidelity. The Stringfield and Teddlie (1991) article gave us some reason for optimism that if the schools were willing to set very high goals, and if they were able to implement the effectiveness fields' findings with high reliability, large achievement gains were possible. Two possible sets of tools were likely to enhance implementation reliability. The first set was the characteristics of HROs.[9]

All three members of the development/research team had participated in previous studies in which well-intended change efforts had failed to achieve strong implementations or desired outcomes. Believing that reliability would only be possible in the context of strong local buy-in, we adopted Datnow's conception of co-construction (see also Berman & McLaughlin's [1978] "mutual adaptation"). If teachers and school heads (principals) co-constructed the reforms, we reasoned that their ownership would be higher, as would our overall chances for success. We explicitly stated that any school that didn't want to work hard at co-creating the reform should not participate, and that we knew a good amount about the various "effectiveness" research bases and HRO principles. However, we always paired those statements with a declaration that the world's leading experts on the specifics of the schools were the local teachers and heads. Success required a melding of expertise.

Groups of schools from three British LEAs asked to participate in the HRS project. We present our data from the Welsh cohort of schools, as it received the most developed version of the reform and as such, presents the most straightforward case description.

Neath-Port Talbot Local Authority: A very successful HRS project

The Neath-Port Talbot (NPT) area is located along the southern edge of Wales, with the Severn Channel as its southern boundary. NPT is about one hour's drive due west of the Welsh capitol of Cardiff. Traditionally, the core of the area's economy was a combination of mining and steel mills. However, the mines were closed over 20 years ago, and the one remaining steel mill is a fraction of its former self. In terms of economic deprivation (poverty), the Neath-Port Talbot area ranks 19th of 22 Welsh districts. That standing has been stable for well over a decade.

The High Reliability Schools project began in Neath-Port Talbot after Professor Reynolds made a presentation to the Welsh Secondary Heads Association. Four heads, three from NPT and one from a neighboring authority, became known as "The Old Welsh Four," and they quickly became advocates of the project. Within a few months, all 11 secondary schools in the NPT LEA had been welcomed in, and they worked as a unified group. The intervention began in the spring of 1996 and continued for nearly four years. The Welsh agreement among the schools and the researchers was as follows:

[8] In retrospect, it is clear that we viewed the HRO research base through the lens of teacher- and school-effectiveness researchers. There are other ways to conceptualize the HRO field (e.g., Weick & Sutcliffe, 2007), but we believe that in the context of education, the principles generally hold up.

[9] For more detailed discussion of this conception of HRO characteristics, see Stringfield (1995); Stringfield, Reynolds, & Schaffer (2008).

- All of the schools would focus on 2–4 very ambitious goals. One required goal was a substantial 5-year rise in the percentage of students obtaining 5 or more A* to C grades on the General Certificate of Secondary Education (GCSE) tests.[10] A second was improved attendance. Each school chose up to two additional goals.
- The heads (principals) would lead the efforts, and the heads and faculties would implement the HRS program schoolwide from the start of the project.
- All schools and departments within schools would agree to share successes and failures, and thus create learning communities within and across schools and LEAs. Each school and department would commit to studying "best practice," both from the international research bases and within and without the HRS schools in England and Wales. The researchers would present school-level series of workshops on the theoretical underpinnings of "High Reliability Organizations," and the research bases on school effects, school change, and teacher effectiveness. Armed with this knowledge, teachers would engage in within- and between-school classroom observations and "no-fault" feedback to peers. Importantly, all agreed that there would be no one piece of research or observational learning required of any school or teacher. HRS was to rely on the well-informed and supported professional judgment of practicing educators in the diverse schools.
- The researchers and administrators of each school would support the faculties in becoming uniquely "data rich." Students would be given short tests as they entered the schools, and age/grade-level teams of teachers would meet and discuss how best to address each student's needs and how to maximize each student's chances of academic success. The resulting student-level data sets were to be (a) rich in individual students' academic histories, (b) available to all teachers and administrators, and (c) regularly shared and discussed by all grade-level teams within schools.
- Almost all of the schools purchased a university-based system of storing and reporting initial intake and eventual GCSE scores. The system made it relatively easy for school personnel to compute a "value added" measure.
- All faculties and administrations committed to regularly review their organization and processes to create widely understood, time-saving Standard Operating Procedures, and to identify and intervene in schoolwide fashion with their pupils who appeared to be at risk of failure.
- A focus on teacher effects/peer observations began immediately. This included both professional development time to learn core aspects of the teacher effectiveness research field (e.g., Brophy & Good, 1986), and for observation in classes within and among schools.
- A strong "departmental effectiveness" component that facilitated within-school learning was emphasized as the project developed.

Several additional components were added over time:

- When the assessment of incoming 11-year-old students at some of the schools indicated many were entering secondary school more than two years behind in reading, an immediate effort was launched to coordinate the secondary school's literacy programs with those of the feeder primary schools.
- The LEA appointed a part-time "HRS Driver" to formally coordinate activities among the Welsh district's schools. The effect was to have HRS continuously "on the radar screen" at each school and in most departments of all schools.
- In Wales, in addition to the "broad brush" principles of HRS and the detailed organizational features of the HRS model as outlined in the components material, there was an additional focus on what came to be called "the little things that matter." HRS meetings increasingly centered upon regular sessions in which each school explained to the whole group of Welsh schools the practical things that they had done at the "micro" level to embed the concepts and the components in the form of practical organizational features at the point of delivery of education to pupils.
- Additional time for professional development was built into the Welsh implementation. Heads and faculties attended regionally based residential sessions (two-day meetings at a conference center) for all head teachers and HRS representatives, and also added

[10] Virtually every British student sits for the GCSEs. The traditional measure of strong academic performance for a student is obtaining "5 or more A* to C" grades on the various examinations (literature, mathematics, various sciences, etc.). Although scores have risen over the last decade, under half of Welsh students obtained 5 or more A*–C grades in the mid 1990s, and those percentages have risen to over half in the first decade of the 2000s.

national residential sessions, all aimed at enhancing knowledge transfers across schools and LEAs. The Professional Development focus in Wales tilted strongly towards turning schools into "knowledge generators" rather than passive knowledge recipients. Particularly, the HRS project focused on introducing peer observation systems to permit the charting, generation, and transmission of good practice in classrooms, training some school personnel to use observation systems that were then cascaded around the entire school.

- The team's focus upon improving schools' capacity to be reflective about their organizational functioning and outputs was enhanced, using additional training. Examples included the provision of sessions on the statistical analysis of data and the provision of a sophisticated, relational database that teachers could access to more efficiently analyse stored grades, background information, and test scores of pupils.
- The program began to take a close interest in the effectiveness of the primary feeder schools that were generating intakes of pupils that, in the case of most schools, were regarded as unintentionally setting "low ceilings" on what it was possible to achieve. Primary senior management teams were invited to the secondary schools' HRS training days. One secondary school went so far as to use some of its own resources to provide a literacy coordinator to the primary schools whose students it served.
- Finally, the HRS representatives and principals received additional materials, some of which focused on topics around being effective managers of change. Additionally, bodies of knowledge to be shared with teachers were first previewed with the head teachers and HRS coordinators, allowing the leadership to be prepared to answer staff questions and ease the material into schools. Heads also selected among possible staff development alternatives based on their perceived needs of faculty and previous efforts that had been successful.

The Welsh GCSE results

The GCSEs are viewed in Great Britain as relatively high-stakes assessments, similar to Advance Placement courses in the United States. For students, a certain number of passing grades are required for such career options as becoming a policeman or postman, and a (higher) number is required for admission to various colleges. Given that all English and Welsh secondary schools have essentially open admissions (i.e., students are not bound to attend the school in their specific geographic area), a rising or falling standing on the percentages of students passing 5+ GCSEs can affect the number of students choosing to attend a school. In turn, teachers' and administrators' positions can be gained or lost. At the low end, the LEA can close secondary schools that have a persistent pattern of very low scores. Well short of that extreme, it is not uncommon for a head teacher to lose his/her job if school-level GCSE scores fall for several consecutive years.

As seen in Table 1, in the three years prior to the HRS project, 16-year-olds in the NPT LEA had scored well below the Welsh national average on the GCSEs. With one exception (discussed below) NPT's modest standing had not raised any hue and cry at the local or national levels. NPT was a relatively deprived area, and the expectations for student performance were modest.

By the end of the intervention, NPT scores had risen essentially to the national average (48.5% vs. 49%). The gain was impressive, but being at the national average did not attract great attention. By the research team's follow-up in 2007, NPT's students were scoring at well above the national average (60.7% vs. 54.2%), and a nationally publicized "value added" assessment had found NPT to be by far the most "value added"

Table 1

Neath-Port Talbot Local Authority, two specific schools, and Welsh national mean percentages of 15–16-year-old students obtaining 5 or more A–C scores on the GCSEs, 1994–2007*

Time Frame/ Group	1994–1996 (pre-)	2000 (post-)	Initial Gain (pre- to 2000)	2007 (follow-up)	Longitudinal Gain (pre- to 2007)
NPT LEA	33.3%	48.5%	15.2%	60.7%	27.4%
Sandfields	14%	35%	21%	47%	33%
Cwmtawe	31%	51%	20%	75%	44%
Wales	40.7%	49%	8.3%	54.2%	13.5%

LEA in Wales (Stringfield, Reynolds, & Schaffer, 2008).

The two schools highlighted in Table 1 tell interesting longitudinal stories. Sandfields secondary is located in a very disadvantaged public housing complex. Because the school's GCSEs had been especially low for several years, the national government threatened to close it. The school head had invited Reynolds to make a presentation to the faculty to determine if there was sufficient interest in participation in the HRS project. Among the questions asked by the faculty was, "If we were to participate, where would you suggest we start?" The school facilities and grounds were in poor shape, and Reynolds suggested starting with a cleanup campaign. The faculty involved the entire community, and in a few weeks, the school's appearance was significantly improved. This gave the faculty a sense of early accomplishment, energy, and hope to go forward. Initially, the majority of the faculty would have been delighted to have achieved a 25 percent or their students obtaining 5 or more A*–C's on the GCSEs. In one year, fully 50 percent of their students earned that high standard. In 1996, the faculty would have thought achieving at that level was impossible.

The second school, Cwmtawe, is located in a more middle-class community. In the three years prior to participating in the HRS project, their level of student achievement on the GCSEs averaged 31 percent. Although this was not viewed as deeply problematic, the school's teachers—and, in particular, the administration—had higher ambitions. Probably no school embraced the HRS principles more fervently than Cwmtawe.

During the implementation years, the school raised its percentage of students obtaining 5+ A*–C's by 20 percentage points (from 31% to 51%) and for the first time exceeded the national average. By 2000, the head, deputies, and teachers had become expert at examining each student's incoming grades and test scores, and at working with the students and their families to produce multi-year plans for each student's success. The result has been that Cwmtawe's scores were well above the national average on 5+ A*–C's, and over the last decade their rate of improvement has been three times the national average. Just as impressive, the school has committed to having all students achieve passing scores on at least some of the GCSEs (in 2007, 98% of students achieved 5+ A*–C scores), and they have set a new goal of having many of their top students obtain 10+ A*–C's. So the school is focusing not just on the state-defined measure, but on high levels of success for all students set by the school. Not every school in the LEA has experienced this level of success, but it is noteworthy that 10 of the 11 secondaries in NPT produced 11-year gains that exceed the national average.

In summary, the Welsh LEA was the third to join the HRS effort and received the more nearly polished presentation throughout. The LEA provided consistent levels of support to its schools, and the heads and teachers were, on average, enthusiastic co-constructors of the reform in their community. The heads took charge of the project from the beginning and probably shared more of their frustrations and successes within and across schools than either of the other groups of schools. The results of this union of researchers—who brought to the table valid findings from previous research—and enthusiastic, improvement-focused local educators continue to speak for themselves.

The English "leafy suburb" district: Our unsuccessful HRS pilot project

In some ways the most complex, and in several ways the least encouraging, of our British HRS LEA stories comes from the first district to suggest and implement the project. Several things went well in this LEA, but several others were problematic. Studying past failures to avoid future ones is a key process in HROs, and we discuss these as part of a prelude to discussing when and where HRS can and probably cannot help local educators improve their schools. Among the strengths in this LEA were several enthusiastic central office staff members, many fine educators in the schools, and enthusiasm inherent in being the first to try to develop a reform. Yet, the project faced several challenges that proved fatal.

The HRS program was developed "on the fly," and there were obvious rough edges on the professional development components, which certainly harmed the project's credibility.[11] The LEA had endorsed two separate reform efforts; eight of the 16 secondary schools in this LEA chose one reform, and the other half chose the other. In the end, neither was successful, and this lack of coordinated focus may have been a substantial problem for both reform efforts. In this initial implementation, we reasoned that the school heads already had very demanding jobs and suggested making deputy heads the "HRS drivers" of the schools. This inadvertently communicated that

[11] Bob Slavin, co-developer of Success for All (SFA), has observed that the first schools to implement almost all new SFA components have been among the least successful. His explanation (personal communication) has been that the rough edges reduce teacher and administrator confidence and commitment.

HRS was of secondary importance in the schools. The effort was initiated by central office staff without initial enthusiasm from individual schools. By contrast, in NPT, local heads had enthusiastically lobbied for HRS buy-in. Finally, at the time of reform implementation, the LEA was already achieving at the national average for percentages of students obtaining 5+ A*–C's, and there was no strong motivator for the schools to take on HRS, or any other demanding reform effort.

Lessons learned

Studies across a range of countries have found that producing measurable change in student achievement is more likely in elementary school reforms than secondary, yet the HRS project produced dramatic student achievement gains in secondary schools. We, as the development team, contributed a full measure to the lack of success in the initial "leafy suburban" site, but the lessons learned from it made contributions elsewhere. One clear lesson is that simply joining an HRS project isn't an automatic route to academic improvement.

Two U.S. educational improvement efforts involving HRO components

Here, we relate two efforts to use HRO principles to enhance reform efforts in the United States.

Grant County, Kentucky

Grant County is a small, rural district in North Central Kentucky, southwest of Cincinnati, Ohio, and northeast of Louisville, Kentucky. Mike Hibbert, Grant County's superintendent, heard about the British HRS project at a conference, and determined to use HRO principles to solve one of his district's more enduring problems—its high dropout rate. An analysis of their data indicated that students who dropped out were unusually likely to have repeated 9th grade.

Mr. Hibbert guided his small central office staff and leaders from Grant County's one middle school and one high school in an effort to focus on helping students succeed in their 8th to 9th grade transition, and, as a result, be more likely to succeed in 9th grade and graduate from high school. The district engaged in an extensive review of issues related to secondary school success and eventually focused on the issue of middle-to-high-school transition. They identified a substantial body of literature on necessary steps for successful transitions (Allen, Christian, & Hibbert, 2010; Morgan & Hertzog, 2001; Oakes, 2009).

A large team of middle- and high-school teachers and administrators took a range of steps, including enhanced student and parent involvement, teacher "intervisitations" between the schools, and 8th-grade student days spent at the high school, all carried out using HRO principles and processes. The results were immediate and dramatic. Hibbert (2010, personal communication) reported that whereas freshman retention had been 38 percent pre-intervention, in the 2009–2010 school year, "the retention rate was 1 percent."

Kentucky mandates and funds all high school students taking the ACT's "PLAN" (pre-ACT) test in 10th grade. In the second year of transition implementation (when the first implementation cohort reached 10th grade and took the PLAN), Grant county's PLAN scores rose a full point (equivalent to a rise of 50 points on the SAT). Further, Hibbert reports that the number of failing grades are down significantly among 9th grade students at his high school.

The Effective Schools for the 21st Century (ES-21) Project

In 2004, the Olin Foundation funded an effort to conduct a "gold standard" (random assignment) study of the implementation of "Effective Schools" variables and processes (Taylor & Bullard, 1995; Teddlie & Reynolds, 2000) in a sufficient number of schools across several states so as to demonstrate the continuing validity of those principles at the standards required by the federal "What Works Clearinghouse." The implementation team decided to use HRS principles as over-arching principles of implementation in the project. The project confronted a continuing series of complications, including these:

- Many LEAs were interested in participating, but attempts to ensure random assignment resulted in over a dozen districts declining offers to participate, even though the project was 100 percent externally funded.
- The requirement of random assignment after they agreed to participate left many principals and teachers feeling that the project was more someone else's research than their own reform.
- No district had as many as 50 percent of its elementary schools participating in the "experimental" group. In several instances, central office staff "borrowed" key ES-21 principles and presented and/or implemented them in control sites.
- Without all schools from a district participating in the experimental component, several superintendents became less than enthusiastic about the idea of some of their schools getting something that others were not.
- The experimental LEAs and schools experienced high rates of professional staff instability, such that re-training became a norm in the project. Four of five LEAs experienced at least one

superintendent turnover; the 17 experimental schools had a total of 35 principals over the three years; some school leadership teams experienced between 100 percent and 200 percent turnover.

Qualitative follow-up interviews consistently found teacher and principal enthusiasm for the project, with educators regularly reporting that they were better prepared to deal with future changes as a result of ES-21 participation. However, in the end, the project did not produce achievement test score gains for the two carefully followed cohorts of students in the study.

Discussion

It is hardly surprising that most school reforms fail. Complex systems—such as schools and school systems—that are inadequately understood and modeled are unlikely to be successful hosts for reforms of almost any type. We assert that research on very complex systems that must succeed in their core missions the first time every time (e.g., High Reliability Organizations) offers guidance for school reform. Efforts to use HRO principles to guide reform in several contexts offer both hope and cautions for future educational reformers:

1. Dramatic improvements in student outcomes are possible, and possible at scale.
2. To achieve those results, the reform components must themselves be based on substantial bodies of research that have demonstrated their value in improving student performance.
3. In educational reform, as in research, "reliability sets the upper boundary of measured validity." HRO research can play a critical role in producing Highly Reliable Schools.
4. Among the conditions necessary for HRSs to evolve are the following:

- As a first condition, both the public and the professional educators must realize that in the 21st century, the costs of educational failure are catastrophic for the individual students who do not achieve their full potential and for the rest of us in society. This is a dramatic shift from 50 years ago.
- Buy-in from both the LEAs and the schools' leadership to a focused set of goals is a critical next step.
- There must be a perception existing, or created early on, that failure to achieve core goals is unacceptable.
- An understanding of—and openness to—the idea that any reform, including HRO-based reform, is a combination of external ideas and continuously evolving local contexts. Just as there are no two air traffic control towers that are alike, there are no two schools needing exactly the same reforms, the same Standard Operating Procedures, and so on. Further, any one school's need for any one SOP may change over time. Dynamic organizations must be dealt with dynamically.
- A minimal level of leadership stability, combined with carefully targeted leadership transitions, is necessary to sustain reliability.
- The HRO characteristics described earlier in this chapter must be followed in detail.

Viewed from an HRS perspective, the conditions that predict reform failure also are knowable. They include, but are not limited to, the following:

1. A lack of initial buy-in to the idea that dramatic improvement in student outcomes is possible.
2. Too many diverse goals.
3. Attempting to implement reforms that are not clearly informed by rigorous research.
4. Lack of multi-year commitment to intensive, shared professional development.
5. Leadership and staff instability, especially if not accompanied with careful, real-time induction into HRS principles.

Conclusions

Complex systems that are inadequately understood and modeled, such as schools and school systems, are unlikely to be successful hosts for reforms of almost any type. It is hardly surprising that most school reforms fail. In this chapter, we have argued that research on very complex systems that must succeed in their core missions the first time, every time (High Reliability Organizations) offers guidance for school reform.

Our first overarching conclusion is that the conditions now exist in which substantially higher educational reliability in the United States is possible. The costs of failure—both for the individual and the society—have become too great for unreliability to continue. Hence, we believe that the country's fundamental choice is not whether to become more reliable, but whether to stumble forward, feeling our way and making many, many mistakes; or whether to understand and control a more efficient process of increasing educational reliability.

The second conclusion allows for a good deal of optimism. Our data indicate that, under specifiable conditions, High Reliability Organization principles can be productively applied in school and district contexts. Thoughtfully, consistently applying HRS principles has produced dramatic results in the United Kingdom and the United States, and could do so again in other schools and districts.

HRS models may initially appear to have a "mechanistic" feel to them. Nothing could be further from the truth. In many schools and indeed countries, schools and systems are "tight" on the processes that are meant to exist and "loose" on the systems to achieve these processes. HRS is the opposite—loose on the precise organizational processes needed, leaving those to be determined in detail by local educational professionals. Where HRS is tight is on specifying the concepts and systems that schools should use to generate their often-different processes.

In an educational world where school systems too often tell their teachers what to do, the HRS model is representative of a different philosophy which sets schools free to determine which research- and proven-practice-based practices to implement. This is surely the way to create a more informed, effective, better supported, and more reliably successful educational profession.

References

Allen, D., Christian, C., & Hibbert, M. (2010). *Lost in transition: A grade 9 transition program using articulation activities.* Louisville, KY: University of Louisville.

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations.* Santa Monica, CA: Rand.

Brophy, J. E., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 328–376). New York, NY: Macmillan.

Collins, J. & Porras, J. (1996.) Building your company's vision. *Harvard Business Review, 74*(5), 65–77.

Datnow, A., Lasky, S., Stringfield, S., & Teddlie, C. (2006). *Systemic integration for effective reform in racially and linguistically diverse contexts.* New York, NY: Cambridge University Press.

Datnow, A., & Stringfield, S. (2000). Working together for reliable school reform. *Journal of Education for Students Placed At Risk, 5*(1&2), 183–204.

Good, T., & Brophy, J. (1987). *Looking in classrooms* (4th ed.). New York, NY: Harper & Row.

Good, T., & Brophy, J. (2007). *Looking in classrooms* (10th ed.). New York, NY: Harper & Row.

Goodlad, J. (1984). *A place called school.* New York, NY: McGraw-Hill.

Heckman, J. J., & LaFontaine, P. A. (2010). The American high school graduation rate: Trends and levels. *The Review of Economics and Statistics, 92*(2), 244–262.

Hibbert, M. (2010, October 24). Transition data for Grant County schools. Personal communication.

Land, D., & Stringfield, S. (2005). Educational governance reforms: The uncertain role of local school boards in the United States. In N. Bascia, A. Cumming, A. Datnow, K. Leithwood, & D. Livingstone (Eds.), *International handbook of educational policy* (pp. 260–280). New York, NY: Kluwer.

Lipsky, M. (1980). *Street-Level Bureaucracy.* New York, NY: Russell Sage Foundation.

Morgan, L., & Hertzog, C. (2001). Designing comprehensive transition plans. *Principal Leadership, 1*(7), 10–18.

No Child Left Behind (2001). Public Law 107-110. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Nunnery, J. A. (1998). Reform ideology and the locus of development problem in educational restructuring. *Education and Urban Society, 30*(3), 277–295.

Oakes, A., & Waite, W. (2009, May). Middle-to-high school transition: Practical strategies to consider. *The Center for Comprehensive School Reform and Improvement Newsletter.* Retrieved from http://www.centerforcsri.org/index.php?option=com_content&task=view&id=669&Itemid=5

Pfeiffer, J. (1989). The secret of life at the limits: Cogs become big wheels. *Smithsonian, 20* (4), 38–48.

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, E. (2002). *World class schools: International perspectives on school effectiveness.* New York, NY: Routledge/Falmer.

Stringfield, S. (1991, January). *A model of elementary schooling effects: Evidence of support.* Paper presented at the meeting of the International Congress for School Effectiveness and Improvement, Cardiff, Wales.

Stringfield, S. (1995). Attempts to enhance students' learning: A search for valid programs and highly reliable implementation techniques. *School Effectiveness and School Improvement, 6*(1), 67–96.

Stringfield, S. (2008). School boards and raising student outcomes: Reflections (Confessions?) of a former urban school board member. In T. Alsbury (Ed.). *The future of school board governance:*

Relevancy and revelation. Lanham, MD: R & L Education.

Stringfield, S., Reynolds, D., & Schaffer, E. (2008). Improving secondary students' academic achievement through a focus on reform reliability: 4- and 9-year findings from the High Reliability Schools project. *School Effectiveness and School Improvement, 19*(4), 409–428.

Stringfield, S., Schaffer, E., Lasky, S., Overman, L., Chrispeels, J., Burke, P., Daly, A., Bolivar, J., Murga, S., & Johnson, P. (2010). *The four quartets: Effective Schools for the 21st Century Project final report*. Louisville, KY: University of Louisville.

Stringfield, S., & Teddlie, C. (1991). Observers as predictors of schools' multi-year outlier status. *Elementary School Journal, 91*(4), 357–376.

Stringfield, S., & Yakimowski, M. (2005). The promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore City Public Schools. *American Educational Research Journal, 42*(1) 43–76.

Taylor, B. (2002, January). The effective schools process: Alive and well. *Phi Delta Kappan,* 375–378.

Taylor, B., & Bullard, P. (1995). *The revolution revisited: Effective schools and systemic reform*. Bloomington, IN: Phi Delta Kappa.

Teddlie, C., & Reynolds, D. (Eds.) (2000). *The international handbook of school effectiveness research*. London, UK: Falmer.

Teddlie, C., & Stringfield, S. (1993). *Schools make a difference*. New York, NY: Teachers College Press.

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*(1), 1–19.

Weick, K., & Sutcliff, K. (2007). *Managing the unexpected* (2nd ed.). San Francisco, CA: Jossey-Bass.

To ask questions, or for more information, write to the authors at their e-mail addresses:
Sam Stringfield: sam.stringfield@louisville.edu
David Reynolds: david@davidreynoldsconsulting.com
Eugene Schaffer: schaffer@umbc.edu

Chapter Three

High Reliability & Leadership for Educational Change

By G. Thomas Bellamy, University of Washington Bothell

Introduction

On April 20, 2010, a deepwater drilling platform used by BP exploded, releasing an 86-day torrent of oil into the Gulf of Mexico and wreaking havoc on the natural environment and on the lives of millions of people, their families, communities, and economies. The aftermath will be felt for generations. When consequences are this catastrophic, the public expects organizations to work without failure, and headlines follow when they do not.

A school that fails may lack the visibility of a BP disaster, but it can have its own catastrophic human consequences for the lives of young people, their communities, and our society. The public understands this all too well, and its high expectations for reliable school performance are reflected in government policymaking and in local news. In these respects, students failing in school might not be all that different from a petroleum company's off-shore accident. As such, the BP disaster might have at least one positive outcome: providing public education with insights about how to avoid failure.

Thanks to research on many accident-free organizations, we know a great deal about managing for success in uncertain situations. These High Reliability Organizations include familiar enterprises such as air traffic control, chemical manufacturing, air travel, electric power generation, and wildland firefighting groups. All face serious hazards in unpredictable circumstances.

HROs achieve reliability through four distinctive organizational accomplishments:[1]

1. Sustaining commitment to a dual bottom line
2. Centralized procedural control and standardization
3. Flexibility for situational improvisation
4. Combining opposite operating modes

Assuming that all structures and strategies are fallible, leaders of successful HROs build the capacity to work in two modes, one standardized and centrally controlled, the other decentralized, improvisational, and situation specific. Each way of operating is always ready to use as shifting circumstances either allow normal work toward organizational goals or threaten reliable performance.

As our understanding of HROs expands, so too does interest in applying their strengths and strategies to manage uncertainty and improve performance in many other organizations, including education (Bellamy, Crawford, Huber-Marshall, & Coulter, 2005; Weick & Sutcliffe, 2007). Stringfield and his colleagues (Stringfield & Datnow, 2002; Stringfield, Reynolds, & Schaffer, 2010; Stringfield, & Yakimowski-Srebnick, 2005) already have shown that HRO strategies, used in combination with concepts from the effective schools literature, can support dramatic and sustained improvements in school learning outcomes.

Of course, public schools are quite different from most HROs. Educators work in public organizations that are naturally open to outside influences; their work is people-and relationship-intensive and depends on far less prescriptive knowledge. Thus, while high reliability seems important to leadership for educational change, contextual differences mitigate against uncritical transfer of HRO practices to schools.

1. Sustaining commitment to a dual bottom line

Successful HROs find ways to balance simultaneous commitments to achieving desired results while avoiding accidents or failure. They are adept at "finding a balance between

[1] Similar to the research on effective schools, qualitative studies of HROs have led to several taxonomies of the distinctive characteristics of these organizations, each highlighting slightly different aspects of reliable operations (e.g., see Roberts, 1990; Rochlin, 1993; Weick & Sutcliffe, 2007). The categories used here highlight challenges in applying HRO strategies in schools.

today's profits and tomorrow's potential disasters" (Roberts, Bea, & Bartels, 2001); they "consider reliability as important an outcome as productivity" (Roberts & Libuster, 1993, p. 16); and they are able to "restate goals in the form of mistakes that must never occur" (Weick & Sutcliffe, 2007, p. 151). They are equally committed to providing service and avoiding failure, based on strong agreement about both the definition and value of success and costs of failure (LaPorte, 1996). For HROs, such dual priorities are not just slogans. The underlying values—commitments to what the organization needs to accomplish and what it should never allow to happen—become cornerstones for the organization's culture (Weick, 1987).

News coverage[2] of BP's accident provides some insight into just how difficult it can be to sustain a cultural commitment to this dual bottom line. Confronted with its history of several prior accidents and safety violations, BP's then-CEO began his tenure with the promise that "the company would make safety its number 1 priority" (*NYT*, 2010, July 12)," and maintained that he had been "laser-focused" (*NYT*, 2010, June 18) on safety while leading BP. But reports of priorities on the ground sound quite different: "Taking shortcuts was ingrained in the company's culture, and everyone in the oil business knew it" (*NYT*, 2010, June 18); "BP was developing a reputation as an oil company that took safety risks to save money" (*NYT*, 2010, May 31).

Why such a discrepancy between management's stated priorities and the operating culture? One possibility is the seemingly disconnected communication about goals for productivity and safety. News coverage can be incomplete, of course, but one set of messages seemed to place an absolute priority on safety, while another attested to the company's total commitment to productivity and profitability ("BP tries to Reassure Shareholders," *NYT*, 2010, July 7). There is little to no evidence of shared company understandings about a balanced commitment to safety and productivity and what this balance means for organizational routines and relationships.

Implications for leading school change: "Balanced and sustainable goals"

Balanced goals are just as important for schools as for high-risk industries. Pressure for educational excellence comes from families and local communities, state and federal policies, and the commitments of educators themselves. The desire to benchmark outcomes against the best in the world (Barber & Mourshed, 2009) simply reinforces these aspirations. At the same time, the language of educational improvement reflects pressure to avoid all academic failure, with its emphasis on eliminating achievement gaps, preventing dropouts, setting non-negotiable goals, and so on. Just like off-shore oil drilling, air traffic control, or hospital operations, it is insufficient for schools to point to the success of some students, however impressive, while others experience failure. *High reliability learning*—bringing all students to proficiency regardless of their circumstances and our challenges—has joined high academic achievement as a paired expectation for public school success. In fact, the escalating requirements for adequate yearly progress increasingly make high reliability learning a precondition for schools to provide other educational services, just as safe operation is a prerequisite for continued operation of most HROs. Like HROs, schools achieve the benefits of balanced goals only when these espoused goals are supported in the school's underlying cultural values. Establishing and sustaining goals in public education is a complex process that requires ongoing engagement by those leading change.

State and federal policies offer strong incentives to define school goals in terms of standardized-test scores, but families and communities expect much more. "We want it all" was John Goodlad's (1984) summary of extensive national research on expectations of schools. In addition to academic learning, communities count on public schools to ensure students' safety and well-being; support social, civic, and ethical development; and to help students pursue individual talents and interests. Since family circumstances and children's needs vary, different priorities emerge from the many demands competing for a school's limited time and resources. And, as public institutions, schools cannot simply settle these priority conflicts through administrative fiat. Instead, the priorities for what schools should achieve and avoid are decided through continuing dialogue, in both the internal and external communities of the school district, which is punctuated by school board elections, funding ballots, labor negotiations, and leadership changes.

In this context, educational leaders face contradictory requirements in their efforts to establish and sustain balanced goals. School goals for what to achieve and avoid must be open to change as a result of ongoing, honest dialogue and political decision making. But a school's goals also must be stable enough to provide a

[2] All news quotations are from the *New York Times* and are available as a set at http://topics.nytimes.com/top/reference/timestopics/subjects/o/oil_spills/gulf_of_mexico_2010/index.html.

foundation for the learning goals, performance indicators, and student assessments that guide the details of change management (Marzano & Waters, 2009). The capabilities needed to pursue any set of achievement and avoidance goals are incorporated over time into the organization's structures, staff skills, and culture. The leader's task is to open opportunities for participation and create the framework for productive discussion (Chrislip, 2002).

Of course, even when goals are developed through broad participation and deliberation, school leaders have the challenge of fostering internal coherence. This is challenging because of the sheer number and variety of educational goals and the opportunity costs associated with any particular set of priorities. Credible communication about balanced goals depends on open discussion of the hard questions about tradeoffs when goals for achievement and avoidance conflict.

For example, when a child is having difficulty with an arithmetic concept and needs extra time, where will that time come from? From the science or art lesson? When the teacher spends extra time with the struggling student, do other children who have mastered the concept miss out on whatever accelerated opportunities might otherwise have been available? Does everyone implicitly agree that there are some activities that cannot be displaced by extra instruction in core subjects?

School leaders may be tempted to avoid the issues, or to leave it up to teachers to decide on a case-by-case basis, but lack of clarity about priorities can also mean lack of timely action to respond to early warnings of failure. Credible communication about real priorities and tradeoffs can also help build the will and capacity to act. It is one thing to have general agreement with a set of priorities and another entirely to reach shared commitment to resilience and a belief that staff and leadership will do everything possible to meet the school's goals for achievement and avoidance.

In their application of HRO strategies to educational change, Stringfield, Reynolds, and Schaffer (2010) asked schools to focus on a very small number of critical goals. Here we suggest two additions: (a) that school goals should explicitly address what should never happen as well as what should be achieved, and (b) that leadership for change should include ongoing commitment to conversation and stewardship of the goals, both internally and externally, in order to preserve balance and achieve stability in school priorities.

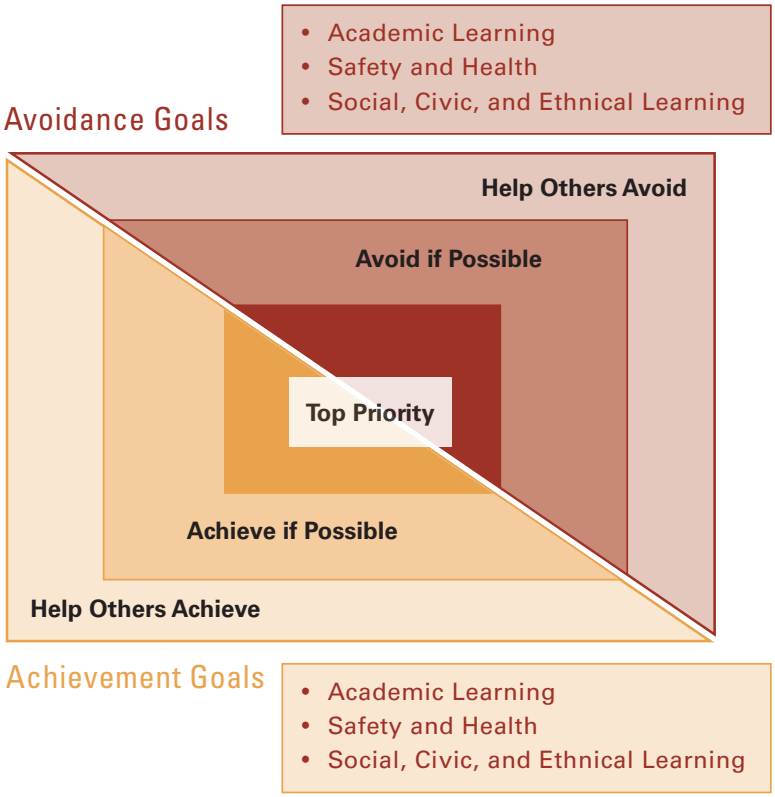
Figure 1 depicts these two additional strategies. The same framework might serve as a way to structure ongoing feedback about how well the school is perceived as implementing those priorities.

2. Centralized procedural control and standardization

Much of the literature on HROs describes organizations with centrally controlled and highly standardized operating procedures. Ways of doing things are designed by experts who can translate the field's best knowledge into practice and supported by management practices, hierarchical decision making, regular feedback, and employee incentives for following those practices. Standard procedures allow the field's current knowledge and the organization's prior learning to be imbedded in routines, provide

Figure 1

Visual prompt to support deliberation about school priorities



a way to coordinate various parts of a complex organization's work, clarify decision authorities and premises, and create a foundation for continuous improvement as processes are evaluated and refined. Successful HROs take advantage of standard operating procedures and central control by continually improving the procedures themselves and the organization's implementation capacity through training, feedback systems, and post-action reviews (Rochlin, 1993).

Although standardization and central control are useful, research on HROs makes it clear that standard procedures are insufficient to achieve high reliability. In HROs, standard procedures are implemented with an assumption of fallibility, constant attentiveness to what might go wrong, and simultaneous investment in the capacity to respond differently when the inevitable problems arise.

Press coverage of the BP accident highlights two challenges related to procedural standardization. The first is whether BP's procedures actually incorporated the best current knowledge or were designed instead to achieve efficiencies by taking greater risks. For example, one report quotes from Congressional correspondence that "some of the decisions appeared to violate industry guidelines and were made despite warnings from BP's own employees" (*NYT*, 2010, June 14). The second challenge is evident in reports that interpreted the failure to standardize procedures as evidence that BP was not learning from previous mistakes. An earlier BP project was described as having "cramped, chaotic conditions…" "It was like having the plumbers, the electricians and the bricklayers come to a construction site at the same time as they are laying the concrete. This was not methodical" (*NYT*, 2010, July 12).

Yet, despite a series of accidents and safety warnings, BP continued to rely on a case-by-case approach, rather than adopt standard operating procedures. While the high reliability literature is more skeptical than the authors of this particular news report about standard procedures as the only strategy to avoid accidents, much of the literature suggests that in most circumstances, standardization is a prerequisite for reliability.

Implications for leading school change: "Skeptical standardization"

The high level of standardization typical of most HROs is seldom present in schools, but the experience of HROs could be more relevant than first appearances suggest. Pressures to achieve high reliability learning are leading an increasing number of schools and districts toward greater central control and standardization in curriculum and instruction. In practice, school leaders standardize normal operations by adopting curriculum materials, pacing guides, common assessments, and so on, which establish consistent instructional practices across a school or district. They also build capacity to implement these standard procedures through staff selection, training, coaching, and evaluation. In such approaches, teachers are expected to comply with and build skills for established programs, communicate with supervisors about problems, and use data to improve program implementation.

Literature on HROs offers two main reasons for procedural standardization. Both seem relevant to public education but application involves special challenges. The first reason is to ensure that the field's best knowledge is applied in any given classroom. Robinson, Hoepa, and Lloyd (2009) call such procedures "smart tools," emphasizing that research knowledge can be built into a set of procedures and routines such that not every user needs to know all of the underlying theory and data in order to receive the benefits. But every teacher knows what an honest critique of research shows: none of our smart tools—no curriculum, program, textbook, or instructional method—works reliably with all students. Thus, there is every reason to argue that standardization alone will not create fail-safe schools. Further, our field's knowledge base contains many different, often contradictory, methods for achieving educational goals (Donmoyer, 1996). And, possibly because different approaches work best with different groups of students, schools are often characterized by intense personal and professional commitments to different curricula and teaching strategies. Conflicts over which program or approach to select are practically assured whenever schools attempt standardization, and such conflict can easily derail any efforts to implement a standardized program. Critics of standardization of instructional procedures also point out that it often causes disruptive shifts during central leadership transitions and is ineffective in supporting teacher learning and development (Levine & Marcus, 2007). Given these limitations, it is reasonable to ask if standardization really makes sense in public education.

This question leads to a second reason for standardization in HROs. From extensive experience consulting with hospitals to reduce medical errors, Resar (2006) observes that, even when more than one research-based approach is available for a given clinical procedure, medical errors increase when each physician continues to use her or his preferred strategy. It's not that one is necessarily better than the others, but rather that, without reasonable standardization, the organization cannot build the

most effective systems—the staff skills, equipment, supplies, scheduling, and other components—that support the clinical procedure. So even when more than one procedure is supported in research, Resar argues that high reliability involves picking one, then building the capacity to use it well.

The circumstances that Resar (2006) describes in hospital care appear analogous to education, where some teacher discretion to choose among research-supported procedures is often expected. Resar's logic suggests, then, that skeptical standardization can be useful as one part of a leadership strategy for school change, as long as efforts to standardize pay special attention to three issues.

- First, standardization makes the most sense when the chosen procedure or program is expected to succeed with a significant majority of students and can be implemented with sufficient slack so that teachers have time to pay special attention to the inevitable group of students for whom the procedure was not successful.
- Second, standardization makes sense when the selected procedures will be used as the foundation for building capabilities in instructional materials, technology, professional development, data systems, and other resources that support instruction.
- Third, precisely because of the limitations of any particular curriculum or procedure, standardization makes the most sense in schools when it is combined with the ability of HROs to shift quickly to in-school improvisation as soon as difficulties arise.

In short, skeptical standardization is most useful when accompanied by its opposite, flexibility for teacher decision making.

3. Flexibility for situational improvisation

A third major accomplishment of HROs is their ability to operate in a flexible, decentralized, and improvisational mode when the need arises. In this mode, critical decisions about what to do are made where the work occurs, not by distant engineers or managers. When seeking expertise in understanding and responding to emerging situations, communication with colleagues replaces vertical reporting (LaPorte, 1996). Knowledge of the situation, combined with expertise to interpret situational developments, replaces general knowledge, standard procedures, and hierarchical authority as the guide to action (Roberts, Yu, & van Stralen, 2004). Operating this way requires open communication, so that those with needed expertise have an opportunity to hear about situations as they develop (Roth, Multer, & Raslear, 2006). To build capabilities for flexible operations, HROs deliberately sustain diverse perspectives and expertise on their staffs, create opportunities for employees to expand professional networks across organizational boundaries, and support norms of resilience that motivate responses to all threats of failure (Weick & Sutcliffe, 2007).

Such flexibility is not freelancing. Responding to situational changes in HROs is a collective process that occurs within constraints of organizational values, collaboration, and previously established decision-making routines (Bigley & Roberts, 2001). The literature contains multiple examples of how otherwise highly controlled organizations also operate in this more flexible mode when needed, including shifts to decentralized and on-site decision making in organizations such as automobile manufacturers (Alder, Goldoftas, & Levine, 1999), aircraft carriers (Roberts, Yu, & van Stralen, 2004), and skyscraper construction (Gawande, 2009).

Implications for leading school change: "Constrained improvisation"

If skeptical standardization enables teachers to use the *profession's accumulated knowledge* in their work, then constrained improvisation[3] enables them to take advantage of *contextual knowledge*—understanding of particular students, families, social groups, and evolving situations that affect learning in a specific classroom and school. In practice, operating in this mode begins with shared goals for student learning and gives teachers the authority to adapt classroom procedures as needed to achieve those goals. School leaders respect the natural differences in instructional approaches that result as teachers respond to current circumstances and implement non-prescriptive strategies for professional development and teacher evaluation.

At first glance, flexibility for situational improvisation seems familiar in schools. After all, behind their classroom doors teachers have traditionally been able to operate as they believed best, while administrators were expected to buffer instructional practice from external pressures. As suggested above, however, HROs use this operating mode in very specific ways that constrain individual flexibility. Improvisation is largely collective work, guided by clear and shared

[3] Originally used in the HRO literature by Bigley and Roberts (2001), the term "constrained improvisation" seems particularly suited to the context of work in public schools, where so many different groups are empowered to influence instructional practice.

commitments to reliable performance, accountability among colleagues, and supported by appropriate checks and balances. In fact, too much workplace discretion in the absence of constraining policies and checks is associated with higher risk of failure (Roberts & Libuster, 1993). Improvisation requires no less coordination and accountability than standardized procedures; it simply achieves these results in different ways.

In schools, collective innovation depends on regular interactions with colleagues, not closed classroom doors. It is based on common goals and involves shared accountability for results. And, like other forms of teacher collaboration, collective innovation often requires administrators to develop new organizational routines that give teachers the time and support to build professional connections with a diverse group of colleagues.

The HRO literature adds an important element to existing literature on teacher collaboration. Much current commentary and procedural recommendations focus on collaboration as a means of teacher learning as a strategy for ongoing improvements in instruction (Levine & Marcus, 2007; Lieberman & Miller, 2008). Constrained improvisation in HROs begins with the more immediate concern of how to respond to a specific situation that threatens reliable student learning. Guided by a commitment to resilience, such collaboration helps to make sense of an immediate situation, develop alternative responses, and make just-in-time adjustments until the threat is addressed. Collaboration for immediate resilient action is not incompatible with longer term teacher learning, nor is it assured when teacher learning is the primary focus.

With these challenges in mind, using constrained improvisation as one part of a strategy for leading school change involves establishing organizational constraints that make that flexibility both collective and accountable while building the capacity to work in a flexible and improvisational mode. To establish the constraints that guide this kind of flexibility, school leaders can do the following:

- Foster shared assumptions and commitments to the school's goals, reliability expectations, commitments to resilience, and considerations for decision making. These elements of school culture can help to coordinate and focus improvisational work without having to rely on a centralized authority structure (Weick, 1987).
- Create school routines and work groups that involve teachers in frequent face-to-face communication about learning challenges. This offers a context for conversations about emerging problems and allows members of a group to bring diverse viewpoints to a collective consideration of planned actions (Gawande, 2009).
- Enhance accountability for improvisational work through post-action reviews, when colleagues can reflect on what was learned from working through difficult situations (LaPorte & Consolini, 1991).

To build the school's capacity to operate in this flexible and improvisational mode, principals can apply three major HRO strategies.

1. Sustain a variety of perspectives and encourage expression of diverse viewpoints.

Schools too often overlook or deliberately discourage diverse approaches in a push to achieve a commitment to an already selected curriculum or instructional approach. Efforts to "get everyone on the same page" may well implement a chosen program, but they will likely weaken the school's ability to operate effectively in an improvisational mode. An immediate challenge for many principals is to be visibly committed to sustaining the knowledge and commitments of those teachers who prefer and are skilled in approaches that are different from the school's adopted programs. It can be tempting to frame these differences as performance problems and pursue transfers or personnel actions. But, these same individuals can contribute important perspectives to the school's improvisational efforts to address inevitable failures in the standard programs.

2. Foster the development of teachers' informal networks and focus communication within these networks on issues of teaching and learning.

Possible approaches include professional learning communities (Dufour & Eaker, 2005), informal networks (Bidwell, 2001), teacher leadership (York-Barr & Duke, 2004), and organizational routines that require regular collective discussion of academic work (Spillane, Mesler, Croegaert, & Sherer, 2009).

3. Establish structures and routines that support collective improvisation whenever needed.

For example, grade-level teams or secondary departments that normally operate to support implementation of standardized programs might also be charged with the responsibility to provide collegial support whenever needed to respond to a student's emerging learning difficulties. Similarly, roles for instructional coaches or district-based curriculum specialists could be designed with dual responsibilities for supporting implementation of standard programs and helping teachers mobilize an early response to impending student failure.

4. Combining opposite operating modes

The two operating modes of HROs are contradictory in many respects. They depend on different organizational structures and routines, contrasting approaches to staffing and training, and different sources of authority for operational decisions. A closer look, however, shows that these two modes also depend on each other. With the pace and complexities of most modern workplaces, employees simply lack the time to devise unique solutions for every circumstance. Some standardization is necessary to provide the slack needed to respond creatively to the most difficult situations. On the other hand, as knowledge becomes more complex, few procedures work in all situations, and even when it seems possible to anticipate every contingency, the result is a system so complex that it can lead to implementation errors that require situational improvisation (Perrow, 1967; 1984). Not surprisingly, the ability to combine opposite operational modes like standardization and flexibility is closely linked to organizational effectiveness (Cameron, Quinn, DeGraff, & Thakor, 2006).

HROs are distinctive in the way they combine these contrasting operational modes. They continuously develop the capacity to operate in both modes and shift between the two approaches as situations arise to threaten reliable performance. HROs use standardized procedures for most normal operations associated with achieving expected results (goals for achievement). Then, operations shift to give situational flexibility when the organization's reliable performance is threatened (goals for avoidance), so that situational sense-making and decisions about action are made where the threatening situation exists. HROs accomplish this shift from standard to flexible operating modes by constantly searching for procedural flaws and situations in which standard routines are unlikely to work. HROs create conditions in which employees notice and communicate about early-stage problems that could threaten reliability and establish incentives for reporting difficulties, even when they might have been caused by the person reporting the difficulty (Roberts, Yu, & van Stralen, 2004). They foster communication channels that allow information about emerging risks to be shared quickly and widely, have shared understandings about when a shift away from standard operating procedures is appropriate, and have ready-to-use structures and routines that help to coordinate the work when shifts are made to the more flexible operating mode (Bigley & Roberts, 2001). Individual employees support this rapid identification by attending carefully to operational details, describing anomalies within an informal network of peers, listening in on others' concerns, and soliciting alternate viewpoints when issues arise (Barton & Sutcliffe, 2009).

Returning to BP's response to the oil spill, news coverage of evolving events highlight the difficulties associated with such shifts between operating modes. The reports prompt one to wonder first about the skepticism that builds attentiveness to emerging problems ("Nobody believed there was going to be a safety issue…", *NYT*, 2010, May 29). Then, as problems became apparent, one asks what would have been required to stop the momentum of daily work toward deadlines long enough to consider alternatives. BP's employees apparently noticed many problems as events cascaded toward disaster ("Documents Show Early Worries about Safety of Rig," *NYT* headline, 2010, May 29; "BP Ignored the Omens of Disaster," *NYT* headline, 2010, June 18). But news reports suggest that this information was never interpreted in ways that prompted a shift to on-site authority for problem solving.

Implications for leading school change: "Public warnings and orderly transitions"

Several ways of combining skeptical standardization and constrained improvisation already exist in public education. In some districts, the boundary between central control and local flexibility is simply the ragged edge of cumulative labor negotiations about whether administrators or teachers should control various decisions. In others, the remnants of organizing schools as loosely coupled systems are evident in the use of central control in some visible aspects of schooling while protecting flexibility in internal classroom operations. More recently, one frequently finds tightly standardized strategies applied to subjects that are included in a state's accountability system combined with more flexibility for teachers to exercise discretion in other subjects (Spillane et al., 2009). As an alternative to using either one approach or the other for various functions, HROs offer the possibility that a school could take advantage of the strengths of both approaches in all of its operations. As with other characteristics of HROs, however, application in education presents unique challenges.

Public warnings

Schools are rich with information about early-stage learning problems, even without waiting for the results to show up in formal data systems. Most teachers already know long before formal testing which students are advancing too slowly to meet expectations, or are being held back by the pace of instruction. But schools are seldom much better than BP at sharing this information, making collective sense

of it, and empowering teachers to work together on solutions. The first leadership challenge, then, is to overcome norms of privacy and autonomy that can limit this communication about emerging learning problems. For example, a school might agree on a shared commitment that no child would fall more than one or two weeks behind peers without collegial discussion of alternatives, so that schoolwide expertise could be tapped quickly in a search for alternative procedures.

A second challenge to effective public warnings results from the structure of specialized programs and professional roles, which can lead school staff to label emerging problems as characteristics of children, rather than results of school procedures. As teachers make sense of learning difficulties, it is easy to jump to the categories for which funding exists, and then hope that specialists can solve the student's problem. Naturally, this chain of responses can limit the range of perspectives about an emerging situation and reduce on-the-spot experimentation and adaptation.

Perhaps the most difficult challenge to an effective early warning system in schools is the *frequency* with which emerging learning problems are encountered. In much of the HRO literature, problems that threaten the organization's reliability are infrequent events. Unexpected fires break out or escape their boundaries, or rare equipment failures require sudden changes in flight plans, and organizations respond with episodic shifts in operating modes.

In schools, student learning difficulties require no less creativity in responding, but these problems occur daily. It is a rare lesson that engages all students and helps each to develop the intended knowledge and skill. In this context, it is unreasonable to assume that more than a small fraction of early warnings will be shared collegially in the course of normal informal conversation during the school day. Consistently sharing information about many emerging problems requires a more systematic approach. A possibility is suggested in Spillane and his colleagues' (2009) discussion of organizational routines that ensure regular collective attention to a particular aspect of school work. One such routine, for example, could be a weekly expectation that members of a professional learning community or grade-level team discuss approaches that are in use with all students who teachers believe are not on pace for success. Protocols could help to structure these conversations around high-impact topics. Research suggests, for example, that communication about early warnings is more likely to interrupt the momentum of normal operations and lead to shifts in strategy when it is accompanied by a request for alternate viewpoints (Barton & Sutcliffe, 2009).

Leadership for this aspect of school change, then, involves developing routines that facilitate sharing of teachers' knowledge about emerging learning problems, establishing expectations for the kinds of learning difficulties that should become public within the school, and fostering norms of mutual assistance to make sense of emerging problems.

Orderly transitions

Noticing and discussing problems, as difficult as these might be, are only the first step. What does it take, then, to interrupt the momentum of activity in a classroom or school long enough to consider alternatives? And when should such consideration actually result in a shift in operating mode?

Without well-understood guidelines for when it is appropriate to shift from one approach to another, combining standardized and flexible strategies could be a recipe for organizational chaos. Can any teacher make this shift at any time, or do certain conditions have to be met first? In the early stages of problem development, it may not be obvious that major changes are needed, but waiting too long for managerial approval can delay needed changes and complicate recovery.

Rules and routines are clearly needed to clarify when a shift to improvisational operations is appropriate, who can reach that decision, and what decision premises should provide guidance. And, because of the frequency with which learning problems are encountered in schools, these routines must be efficient, allowing the school to address multiple threats to reliability simultaneously while sustaining normal operations for other aspects of the work. In schools, then, shifts in operating modes are an ongoing part of operations, unlike the episodic shifts that are more typical of HROs. In effect, schools need to operate in both modes all the time, as different aspects of the work and different students' challenges require improvisational strategies. The resulting challenge is to keep track of the shifting functions that are being addressed in each mode and communicate these well enough to enable coordination of work, collegial support, and supervision.

Gawande's (2009) analysis of the impact of various types of checklists in improving reliability suggests two further requirements for switching from standardized to flexible operating mode. First, the routine should involve collegial discussion, so that decentralized decision making takes full advantage of the expertise of everyone with relevant knowledge, and, second, that it be accountable, in the sense that progress and results are regularly reviewed.

An ongoing case study of one particularly successful elementary school illustrates how such a routine could operate. Teachers are expected to follow prescribed instructional procedures for normal operations, but they also report that they can begin experimenting with alternatives in order to solve an emerging student-learning problem. They do this, however, *after they have discussed the problem with other members of their grade-level team*. The grade-level team, thus, has two different functions. Most of the team's activity focuses on coordinating the regular work of implementing standard programs. But the established relationships, physical proximity, and shared responsibility also create an efficient context for sharing early concerns about student learning and soliciting advice from peers. The authority to move ahead after this discussion with team members supports more rapid adjustments than would be possible if supervisor approval were required. Also, it adds a measure of collective expertise and peer accountability that would be absent if teachers simply made changes that they felt were needed without consultation.

A theory of action for leading educational change

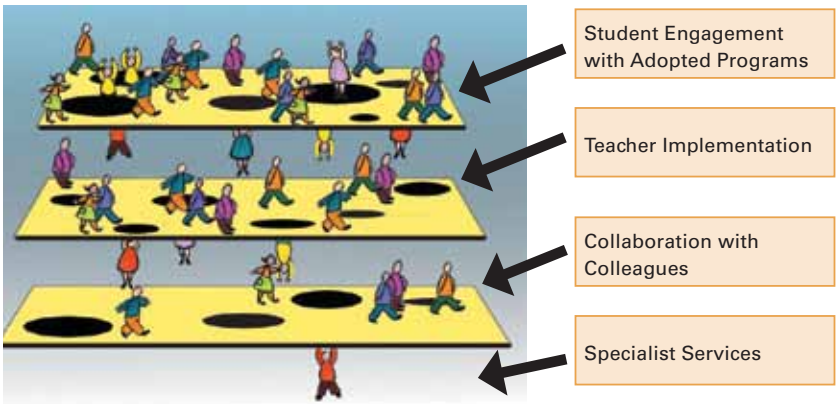
In the preceding discussion, the four distinctive accomplishments of HROs and the corresponding challenges associated with their use in schools combine to frame several strategies for leading educational change. Here is a summary of the larger theory of action for leading school change that these strategies comprise:

- **Balanced and sustainable priorities.** If school leaders establish priorities for what the school should achieve and avoid through inclusive deliberation and open communication about required trade-offs, the resulting goals will be more likely to be sustained long enough to build the capabilities, structures, and supportive cultures that allow the school system to succeed.
- If schools deliberately develop capacities to work simultaneously in **skeptical standardization** and **constrained improvisation** modes, they will be more likely to achieve both high performance and high reliability.
- **Skeptical standardization.** If standardized procedures are used effectively and skeptically, so that they work with most students, allow slack to deal with exceptions, and are subjected to constant watchfulness for inevitable problems, they will provide an important foundation for reliable performance.
- **Constrained improvisation.** If flexibility for innovation is used with constraints that ensure access to collective expertise and shared accountability for results, it will be a powerful means of addressing the wide variety of problems that emerge as any instructional program is implemented.
- **Public warnings.** If the school creates incentives and routines for sharing information about learning problems in the earliest stages, collective expertise can be brought to bear before those problems cascade into intractable failures.
- **Orderly transitions.** If the school has clear routines and decision frameworks for when operations should shift from skeptical standardization to constrained improvisation, then it is more likely that the strengths of both approaches will be used to achieve goals and avoid failures.

Reason (2000) offers a useful metaphor for how the school's core work of teaching and learning could be shaped by this theory of action for leading school change. He compares HROs to several layers of Swiss cheese, each able to prevent some—but not all—problems from slipping through. As long as holes do not line up, organizations that create several "slices" can effectively prevent failures. Thus, successful HROs construct reliable systems out of several protective layers, each unreliable by itself.

Figure 2

A swiss cheese model of high reliability schooling



Note: Adapted from Bellamy, Fulmer, and Muth (2007). Used with permission of the authors.

Similarly, the challenge of change management in education can be seen as one of constructing and sustaining several layers—program components and strategies—that support a school's achievement and avoidance goals. Figure 2 illustrates one possibility. Here, the first layer is the adopted curriculum, together with students' independent engagement with the associated materials. In the present theory of action, this layer is constructed as school leaders develop curriculum frameworks to achieve the school's priority goals and adopt particular curricula and programs as part of a plan for skeptical standardization. The layer is strengthened as school leaders select research-based programs, ensure access to support materials, and communicate about the importance of the learning objectives. For some students, this is sufficient for learning.

The second layer involves the teacher's implementation of the standard programs, including explanations, questioning strategies, pacing, and task assignments as well as the more general classroom routines and relationships with students. School leaders support this layer through such activities as professional development and coaching in use of the adopted programs, data systems that provide regular feedback, and opportunities to observe and learn from colleagues who are using the same programs.

The third layer consists of the added resources that a team of colleagues can bring to the task of making sense of student learning problems and responding adaptively when they arise. This layer represents the shift from skeptical standardization to constrained improvisation and offers the possibility that collective teacher expertise can produce effective interventions for some emerging problems. Supporting this layer means fostering diverse skills and perspectives among teachers, building routines that foster rapid communication about emerging problems, and having clear understandings about when teachers are authorized to shift out of normal procedures.

A final layer involves use of the school's formal programs for students having difficulty, which typically bring additional resources, expertise, and formal procedures to address continuing difficulties. As a theory of action for leading school change, then, strategies from High Reliability Organizations offer a complex but practical strategy for connecting school organization and management with the core work of teaching and learning.

Conclusion

BP's tragic accident in the Gulf of Mexico may be relevant for educational leaders for two reasons. First, it highlights just how complex and difficult operating without failures can be, helping to clarify the distinctive capabilities and accomplishments through which HROs achieve accident-free performance. Each of these accomplishments is a major leadership challenge, and each frames a significant agenda for those responsible for educational change.

The second lesson for educators from the BP experience, more implicit in the preceding discussion, lies in the similarities between reports of the company's operations leading up to the accident and many current conditions in public education. As educators, we confront similar discrepancies between our non-negotiable goals for student learning and the operating cultures of many schools. We often fail to achieve the benefits of standardization with half-hearted implementation of best practices, but then also miss out on the benefits of improvisation due to insufficient support for collaboration. And we allow early warnings of progress, which are clearly evident to some, to slip by without public discussion and collective action, giving emerging problems time to cascade into intractable failures. Addressing these and other threats to high reliability challenges educators to continue to improve strategies for leading school change.

In sum, the accomplishments of HROs offer a general theory of action for leading school change with equal commitment to what schools should achieve and what they should avoid. Reaching these dual goals requires an ability to operate simultaneously in two operating modes, one centrally controlled and standardized, the other distributed and improvisational. Shared information about emerging problems allows these contrasting modes of operation to be combined in orderly ways. Application of these HRO concepts to educational change is still in an early stage of development, and many details still depend on extrapolation from experience in other settings. Nevertheless, promising results from initial school applications and still-unfulfilled expectations for high-reliability learning offer strong encouragement to continue exploring what schools can learn about avoiding failure from High Reliability Organizations.

References

Alder, P., Goldoftas, B., & Levine, D. (1999). Flexibility versus efficiency? A case study of model changeovers in the Toyota production system. *Organization Science, 10*, 43–68.

Barber, M., & Mourshed, M. (2009). *How the world's best performing schools come out on top.* Washington, DC: McKinsey & Company.

Barton, M., & Sutcliffe, K. (2009). Overcoming dysfunctional momentum: Organizational safety

as a social achievement. *Human Relations, 62*(9), 1327–1356.

Bellamy, T., Crawford, L., Huber-Marshall, L., & Coulter, G. (2005). The fail-safe schools challenge: Leadership possibilities from high reliability organizations. *Educational Administration Quarterly, 4*(3), 383–412.

Bellamy, T., Fulmer, C., & Muth, R. (2007). Five ideas for reframing the principalship. *Educational Leadership and Administration, 19*, 57–74.

Bidwell, C. (2001). Analyzing schools as organizations. *Sociology of Education, 2001 Extra Issue,* 100–114.

Bigley, G., & Roberts, K. (2001). The incident command system: High reliability organizing for highly volatile and complex environments. *Academy of Management Journal, 44*(6), 1281–1299.

Cameron, K., Quinn, R., DeGraff, J., & Thakor, A. (2006). *Competing values leadership: Creating value in organizations.* Northampton, MA: Edward Elgar.

Chrislip, D. (2002). *The collaborative leadership fieldbook.* San Francisco, CA: Jossey-Bass.

Donmoyer, R. (1996). The concept of a knowledge base. In F. Murray (Ed.), *The teacher educator's handbook: Building a knowledge base for the preparation of teachers.* San Francisco, CA: Jossey-Bass.

Dufour, R., Eaker, R., & Dufour, R. D. (Eds). (2005). *On common ground: The power of professional learning communities.* Bloomington, IN: National Educational Service.

Gawande, A. (2009). *The checklist manifesto: How to get things right.* New York, NY: Metropolitan Books.

Goodlad, J. (1984). *A place called school.* New York, NY: McGraw Hill.

LaPorte, T. (1996). High reliability organizations: Unlikely, demanding, and at risk. *Journal of Contingencies and Crisis Management, 4*(2), 60–71.

LaPorte, T., & Consolini, P. (1991). Working in practice but not in theory: Theoretical challenges of high reliability organizations. *Journal of Public Administration Research and Theory, 1*, 19–47.

Levine, T., & Marcus, A. (2007). Closing the achievement gap through teacher collaboration: Facilitating multiple trajectories of teacher learning. *Journal of Advanced Academics, 19*(1), 116–138.

Lieberman, A. & Miller, L. (2008). *Teachers in professional communities: Improving teaching and learning.* New York, NY: Teachers College Press.

Marzano, R., & Waters, T. (2009). *District leadership that works: Striking the right balance.* Bloomington, IN: Solution Tree.

Perrow, C. (1967). A framework for the comparative analysis of organizations. *American Sociological Review, 32*, 194–208.

Perrow, C. (1984). *Normal accidents: Living with high risk technologies.* New York, NY: Basic Books.

Reason, J. (2000, March 18). Human error: Models and management. *British Medical Journal, 320*, 768–770.

Resar, R. (2006). Practical applications of reliability theory: Making noncatastrophic health care processes reliable. *Health Services Research 41*(4), 1677–1689.

Roberts, K. (1990). Managing high reliability organizations. *California Management Review, 32*, 101–113.

Roberts, K., Bea, R., & Bartles, D. (2001) Must accidents happen? Lessons from high-reliability organizations. *Academy of Management Review, 15*(3), 1–12.

Roberts, K., & Libuster, C. (1993). From Bophal to banking: Organizational design can mitigate risk. *Organizational Dynamics, 21*, 15–26.

Roberts, K. H., K. Yu, & D. van Stralen (2004). Patient safety as an organizational systems issue: Lessons from a variety of industries. In B. Youngberg and M. Hatlie (Eds.), *Patient safety handbook* (pp. 169–186). Sudbury, MA: Jones and Bartlett.

Robinson, V., Hoepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why.* Auckland, NZ: University of Auckland and New Zealand Ministry of Education.

Rochlin, G. (1993). Defining "high reliability" organizations in practice: A taxonomic prologue. In K. Roberts (Ed.), *New challenges to understanding organizations* (pp. 11–32). New York, NY: Macmillan.

Roth, E., Multer, J., & Raslear, T. (2006). Shared situational awareness as a contributor to high reliability performance in rail yard operations. *Organization Studies, 27*(7), 967–987.

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. In C. Cazden (Ed.), *Review of research in education* (vol. 16, pp. 353–389). Washington, DC: American Educational Research Association.

Spillane, J., Mesler, L., Croegaert, C., & Sherer, J. (2009). *Coupling administrative practice with the technical core and external regulation: The role of organizational routines.* Evanston, IL: Institute for Policy Research.

Stringfield, S., & Datnow, A. (2002). Systemic supports for schools serving students placed at risk. In S. Stringfield & D. Land (Eds.), *Educating at-risk students* (pp. 269–288). Chicago, IL: National Society for the Study of Education.

Stringfield, S., Reynolds, D., & Schaffer, E. (2010). Improving secondary students' academic achievement through a focus on reform reliability. *School Effectiveness and School Improvement, 19*(4), 409–428.

Stringfield, S., & Yakimowski-Srebnick, M. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore City Public Schools. *American Educational Research Journal, 42*(1), 43–75.

Weick, K. E. (1987). Organizational culture as a source of high reliability. *California Management Review, 29*, 112–127.

Weick, K., & Sutcliffe, K. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty* (2nd ed.). San Francisco, CA: Jossey-Bass.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research, 74*(3), 255–316.

Tom Bellamy is professor of education at the University of Washington Bothell. He can be contacted by e-mail at tbellamy@uwb.edu.

Chapter Four

Best in the World:
High Performance with High Reliability

By James H. Eck, McREL

Despite low rankings, a cause for hope

At McREL's Best in the World exploratory gathering (described in Chapter One), Martin West previewed findings from a study conducted by his colleague from Harvard, Paul Peterson, who, along with Eric Hanushek and Ludger Woessmann, compared the mathematics achievement of top-performing students in the United States to that of students in other OECD countries. Once again, the news isn't promising. Using a cross-comparison study calibrating PISA with NAEP results,[1] Hanushek, Peterson, and Woessman (2010) found that no less than 30 of the 56 other countries that participated in the PISA math test had a larger percentage of students who scored at the international equivalent of the advanced level.[2] Twelve other countries had more than twice the percentage of highly accomplished students as the United States (Hanushek, Peterson, & Woessman, 2010).

In addition, shortly after this October gathering, results from the 2009 PISA were released. Out of 34 countries, the United States ranked 14th in reading, 17th in science and 25th in mathematics (OECD, 2010). Those scores, although higher than those from 2003 and 2006, still lag far behind the highest scoring countries, including South Korea, Finland, Singapore, China, and Canada. "This is an absolute wake-up call for America," said U.S. Education Secretary Arne Duncan. "The results are extraordinarily challenging to us and we have to deal with the brutal truth. We have to get much more serious about investing in education" (Armario, 2010). Secretary Duncan's voice joins many others who have suggested we have become a nation of sound sleepers when it comes to educating our children.

As bleak as the PISA data may look for the United States, Sir Michael Barber, in his opening presentation at the Best in the World gathering, expressed cause for hope and shared a quote gained from his days as a student at Oxford: British historian George Trevelyan, in his three-volume biography of Garibaldi, wrote, "There come rare moments, hard to distinguish but fatal to let slip, when all must be set upon a hazard" (Barber, 2010). Barber optimistically observed the stars are lined up in the United States, with the new common standards and a push for common assessments, Race to the Top (even with its flaws), and perhaps most importantly, what he sees as a sense of a national effort to address the achievement gap.

Our focus for the exploratory gathering and this monograph has been expressly on "the other achievement gap" between the United States and other systems of education worldwide. Yet, in light of Schleicher's (2010) conclusion, the variability in achievement gaps found among and within state and district educational systems across the United States demands an equivalent amount of our attention. McKinsey & Company (2009) purport: "In fact, the most striking, poorly understood, and ultimately hopeful fact about the educational achievement gaps in the United States involves the huge differences in performance found between school systems, especially between systems serving similar students" (p. 12).

To once again become among the "best in the world," we at McREL, along with Bellamy and Stringfield, Reynolds, and Schaffer, believe we should not only be looking outward to the highest performing educational systems, but also looking outward to High Reliability Organizations. As noted in Chapter Two, Stringfield, Reynolds, and Schaffer establish

[1] The NAEP scores came from 8th graders in 2005, while PISA 2006 was administered one year later to students at the age of 15, the year at which most American students are in 9th grade.

[2] While just six percent of U.S. students scored at or above the advanced level cut score on the PISA 2006 exam, 28 percent of Taiwanese students did. At least 20 percent of students in Hong Kong, Korea, and Finland were in the advanced category.

a position that the urgency for high reliability evolves from a realization that: (1) failures of the system have catastrophic results, (2) current levels of performance variability are unacceptable, and (3) much higher levels of performance reliability are possible.

Building a foundation for "true" educational HROs

To get clearer on the concepts, structures, and processes evident in HROs, it is important to first identify the core principles and practices of true High Reliability Organizations. Although research and theory-building on accidents, human performance, and high reliability began earlier (Perrow, 1999; Roberts, 1990), we are grounding our theory in the work of Karl Weick and Kathleen Sutcliffe (2001, 2007). The remainder of this chapter synthesizes their research, the ideas from Stringfield, Reynolds, Schaffer, and Bellamy from the previous chapters, and concludes with McREL's thinking about how principles, characteristics, and strategies from HROs translate to educational systems.

Stringfield, Reynolds, and Schaffer suggest a "best in the world" education involves two components: (1) knowing what works extremely well, and (2) providing it with remarkable reliability (p. 1). They approach reliability through the lens of school effectiveness research to establish "what works" with the capacity to more reliably deliver it. Their long history of school effectiveness research, coupled with their 11-year study of HRS systems in the United Kingdom, give them a strong basis for their claim. Their results, particularly in the county of Neath-Port Talbot, Wales, indicate that High Reliability Organization principles can be productively applied in school and district contexts.

With their HRS project, Stringfield, Reynolds, and Schaffer were attempting to merge their abstract understanding of HROs with school effectiveness findings to provide practical guidance to educators. However, in Chapter Three, we saw that Bellamy took a different approach by looking at how schools work through a lens of organizational and change theory. By adding our own research and literature base, McREL is building a theory of action for high-reliability education systems. And our view stems from a perspective gained through conducting several meta-analyses and research syntheses of instruction, school-level effects, extended learning, and school and district leadership. We have most recently synthesized this research into the publication *Changing the Odds for Student Success: What Matters Most* (Goodwin, 2010), in which we present the What Matters Most framework, composed of the following components:

- Guarantee challenging, engaging, and intentional instruction.
- Ensure curricular pathways to success.
- Provide whole-child student supports.
- Create school cultures with high expectations for behavior and learning.
- Develop data-driven, high-reliability systems.

If we know what works, why aren't we doing it? McREL's explanation for this is two-fold: (1) what we know about best practice in teaching and in leadership is not being practiced with superior execution, and (2) our educational systems are not well designed to achieve high performance with high reliability. This is both a people problem and a system problem; some suggest it is *mostly* a system problem. High Reliability Organizations recognize that people will make errors and mistakes, and unless they are intentional, even those should be considered system issues.

Let's return to our working definition of high-reliability educational systems from Chapter One: *high levels of student performance, achieved as a result of high-quality instruction, delivered through superior execution of effective research-based practices, with low variability in the quality of instruction within and between schools.*

HRO principles and characteristics

Weick and Sutcliffe (2001, 2007) outline five principles of High Reliability Organizations. These five have been interchangeably referred to by the authors and others as principles, hallmarks, and dimensions. We refer to them as principles, and although their names have changed slightly over time, we use them as follows:

1. Preoccupation with failure
2. Reluctance to simplify interpretations
3. Sensitivity to operations
4. Commitment to resilience
5. Organizing around expertise

McREL is using these principles as the foundation for developing a theory of action for high-reliability educational systems. In addition to these principles, there are a few key characteristics of HROs that researchers have identified, as well as a variety of structures, processes, and strategies that specifically transfer to educational systems, as you will see later in this chapter.

Principles of High Reliability Organizations

1. Preoccupation with failure

High Reliability Organizations focus on errors and mistakes. This doesn't mean they are paralyzed by anxiety about what could go wrong, or that

they fear personal or organizational failure. High Reliability Organizations do, however, adhere to the slogan coined by NASA during the near-catastrophic Apollo 13 mission that "failure is not an option." There is no acceptable level of loss for a high-reliability organization.

This unwavering attention to the first signs of events that can cascade toward catastrophic failure, or "weak signals" (Weick & Sutcliffe, 2001), positions HROs to respond early and at the source of the problem before it escalates. Just as importantly, HROs do not become complacent with success. Traditional HROs operate continuously under high-risk conditions yet demonstrate safety records approaching 100 percent. Coming close is not acceptable because failure means that lives can be lost.

What if school systems considered student failure as catastrophic as an airplane failing to land safely or a patient failing to recover from surgery? Moreover, what if educators viewed student failure not as the fault of the child, but as a failure of the system? For many, this will require changing core beliefs and assumptions about education. Our standards may have evolved beyond the "sort and select" model of the Industrial Age, but we continue to expect some students to succeed in school and some to fail.

2. Reluctance to simplify interpretations

High Reliability Organizations are highly complex, interconnected systems, technologically and in the amount of human interactions. Humans as a species are very good at finding patterns, but this trait also predisposes us toward categorizing what we observe into what we already know. It subjects us to blind spots where "believing is seeing" (Weick, 2011).

Hoy, Gage, and Tarter (2006) explain that schools need to simplify less and "see" more. Knowing that life in schools is complex, teachers and administrators need to adopt multiple perspectives to understand the shadings that are hidden below the surface of the obvious. While avoiding oversimplification, HROs don't get so lost in complexity that they do not take action. They utilize sophisticated data collection systems and analysis processes to drill down to the root cause of the problem. They do something and evaluate the response within the system.

Some districts focus their sole attention on a *post-mortem* evaluation of performance on state and national standardized assessments. These data may be helpful for comparing schools and districts and even for program evaluation, but the information comes too late and is of little value for identifying individual student difficulties and responding with real-time intervention.

Other, more reliability-oriented systems use a repertoire of assessments and focus especially on diagnostic and progress monitoring measures in a Response to Intervention strategy. Oliver Wendell Holmes, Jr. nicely sums up this principle with the statement: "I would not give a fig for the simplicity this side of complexity, but I would give my life for the simplicity on the other side of complexity."

3. Sensitivity to operations

HROs are attentive to the front line, where the real work gets done. (Weick & Sutcliffe, 2007). Sometimes it is referred to as situational awareness, "having the big picture of the moment" (p. 32). This is facilitated by constant interaction and communication throughout the organization, which includes frequent operations meetings, widely distributed real-time measures of performance, and frequent face-to-face interaction. In schools and school districts, sensitivity to operations may be the guiding principle to drive the effective implementation of professional learning communities (PLCs).

HROs do not allow hierarchies to become dysfunctional bureaucracies (Weick & Sutcliffe, 2001). For Hoy et al. (2006), this principle means staying close to the core function of the organization. For educational systems, the technical core of what we do is teaching and learning. As McKinsey & Company (2007) conclude, it's all about instruction.

Sensitivity to operations is also about empowering highly competent individuals closest to the event with the ability and responsibility to push the button or throw the switch. Anomalies are noted while they are still tractable and can still be isolated (Weick & Sutcliffe, 2007) and are acted upon before they become a full-blown unexpected event.

4. Commitment to resilience

Despite their best efforts at attending to weak signals of impending failure, HROs do experience failures. However, they also construct multiple preventative measures and containment systems to minimize the effects of accidents, anticipating that the unexpected may happen.

HROs recognize it is impossible to avoid human errors altogether (Weick & Sutcliffe, 2001). They develop capabilities to detect, contain, and bounce back from those inevitable errors that are part of an indeterminate world. When the unexpected happens, the organization rebounds with persistence, resilience, and expertise (Hoy, Gage, & Tarter, 2006). Resilience is that characteristic which encourages people to act while

thinking or acting in order to think more clearly (Weick & Sutcliffe, 2007).

5. Organizing around expertise

HROs cultivate diversity of expertise and perspective. Their focus is on matching expertise with the problem regardless of rank or status (Hoy et al., 2006). Rigid hierarchies have increased vulnerability to errors (Weick & Sutcliffe, 2007). Instead, the decision structure in effective HROs is a hybrid of hierarchy and specialization (Weick & Sutcliffe, 2001). Important decisions are made by important decision makers. The twist, according to Weick and Sutcliffe, is that the designation of who is important migrates to the person or team with acknowledged, problem-specific expertise. This is often a dynamic process, where knowledgeable people self-organize into ad hoc networks to provide expert problem solving. In schools, PLCs should flexibly and adaptively use all the human assets available to them.

Acting with anticipation and containing the unexpected

You probably have gathered from the descriptions of the five principles of HROs that they are highly interconnected. Preoccupation with failure, reluctance to simplify interpretations, and sensitivity to operations together establish a set of principles and repertoire of processes that Weick and Sutcliffe (2001) refer to as "acting with anticipation." Simply put, HROs work to anticipate the unexpected and prevent small errors and mistakes from occurring in the first place. A commitment to resilience and cultivating deference to expertise enable HROs to contain the unexpected. The HROs Weick and Sutcliffe studied first tried to build in prevention, and then intentionally avoided becoming so complacent that they had prevented all errors. They tended to adopt organizational mindsets of seeking the early signs of failure and finding remedies quickly.

Beyond the five principles, there are additional characteristics of HROs that deserve attention, particularly because of their potential in developing a theory of action for high reliability in educational systems. One such characteristic is "mindfulness."

Mindfulness

High Reliability Organizations attend to the five principles through a constant state of *mindfulness*. Weick and Sutcliffe (2001, 2007) expand Langer's (1989) conception of individual mindfulness to the level of the organization. HROs are characterized by "an underlying style of mental functioning that is distinguished by continuously updating and a deepening of increasingly plausible interpretations of what the context is, what problems define it, and what remedies it contains" (Weick & Sutcliffe, 2001, p. 3). A mindful organization is more than the sum of mindful individuals (Hoy, 2003).

Mindful organizations also manage the unexpected in early stages, when the signals of trouble are subtle and weak. They encourage the reporting of errors and any failure, no matter how small, as a window to the functioning of the system as a whole (Weick & Sutcliffe, 2001), and develop "a rich awareness of discriminatory detail" (p. 32).

Mindful organizations develop and use enabling structures and processes that enable error identification and correction, cooperation, collaboration, innovation, improvisation, and creativity. Conversely, mindless organizations develop and utilize inhibiting structures and processes. Mindlessness is characterized by "a style of mental functioning in which people follow recipes, impose old categories to classify what they see, act with some rigidity, operate on automatic pilot, and mislabel unfamiliar new contexts as familiar old ones" (Weick & Sutcliffe, 2001, p. 92). Although traditional HROs are hierarchically structured and have tightly coupled processes, they realize the need to be flexible, adaptive, and responsive. Rigid bureaucracies are not conducive to mindfulness; in fact, they may produce a mindless standardization (Hoy, 2003).

A key strategy for encouraging mindfulness is the use of after action reviews (AARs). Senge (2006) calls the Army's AARs "arguably one of the most successful organizational learning methods yet devised." Not to be outdone by the Army, the Navy refers to their process as "during action reviews." Wildland firefighters call their process "lessons-learned reviews." Stringfield, Reynolds, and Schaffer advise that, in order to maintain an ongoing, multi-level alertness to surprises or lapses, HROs build powerful databases that possess relevance to core goals, rich triangulation on key dimensions, and real-time availability. At McREL, we regularly conduct after action reviews, particularly following large projects or events. The purpose of AARs is to learn as an organization, not to place blame or single out individuals. We ask ourselves three questions: (1) What went right and what went wrong? (2) What did we learn? and (3) How can we use this information to improve?

Key elements of high reliability educational systems

We recognize that educational systems are inherently different from those organizations that have traditionally, and accurately, been identified as demonstrating high reliability. It may be a stretch to think of school districts and schools in

terms of "failure-free" operations, and it may be even more of a stretch to put the HRO principles into practice. Nonetheless, under the umbrella of organizational mindfulness, we believe the following key elements from HROs should be in the formula of consistently high-performing educational systems:

- Focus on a few key goals.
- Establish standard operating procedures (SOPs).
- Design structures and processes for defined autonomy and constrained improvisation.
- Create and maintain safe reporting cultures.

Focus on a few key goals

What should be evident by now is that the overarching philosophy of HROs is a preoccupation with failure, translated into goals that everyone in the system not only can articulate, but practice with unwavering attention. Kathleen Roberts (1990), from her research of flight deck operations on nuclear aircraft carriers, relayed this insight from a lower-ranking deckhand: "This is just a bird farm. The birds come in, they get fed, and they go" (p. 172). Stringfield and colleagues in their HRS research, note the importance of defining a clear and finite set of goals, shared at all organizational levels. They stress the need for these goals to be co-constructed between the researchers/reformers with teachers, school leaders, and school systems.

Similarly, McREL's research on district-level leadership highlighted the importance of specifying a few *non-negotiable goals*, at the district level, that should include goals for student achievement and instruction. Just as important (and statistically significant) was the need for these goals to be collaboratively developed. Once the non-negotiable goals are established, a third "responsibility" is alignment with and support of those goals, through all levels of the organization (Marzano & Waters, 2009).

Bellamy refers to these as "balanced and sustainable goals," which constitute the first of his organizational accomplishments of HROs. In Chapter Three, he explained that HROs hold a dual bottom line—balancing commitments to both safety and productivity. These are translated into goals for achievement, coupled with goals for avoidance. Drawing upon McREL's findings, Bellamy advised that school goals be honestly open to change as a result of ongoing dialog and political decision making, while being stable enough to provide a foundation for learning goals, performance indicators, and an instructional program.

Establish standard operating procedures

From McREL's perspective, the highest performing systems in the world establish and accomplish non-negotiable goals for instruction that translate into practice in every classroom. In order to increase the quality of instruction and reduce the variability in instructional quality, they establish clear instructional priorities at the system level, establish a systematic and systemwide approach to instruction, invest in teacher preparation and professional development, and develop strong instructional leadership. In other words, they very carefully develop tighter coupling within the system for curriculum and instruction.

If we know what works from decades of effective teaching and effective schools research, in terms of research-based best instruction, we must ask ourselves, "Then why aren't we doing those things consistently?" By suggesting standard operating procedures for instruction, we are not suggesting lock-step adherence to a particular instructional approach. Some districts, it seems, have gone too far with their implementation of curriculum pacing guides to the point where every teacher is expected to be on the same page on the same day. In *Classroom Instruction that Works: Research-based Strategies for Increasing Student Achievement* (Marzano, Pickering, & Pollock, 2001), the authors identified nine categories of instructional strategies that correlate with high levels of student achievement. However, mindlessly employing the strategies will not raise student achievement; teachers must understand and act on how, when, and why to use them.

Stringfield et al. suggest that regularly repeated tasks that are determined to be effective should become Standard Operating Procedures. These SOPs do not only include instructional strategies identified from the effective schools research, but also time-saving efficiency measures and identification/intervention procedures for students who appear at risk of failure. Stringfield et al. are quick to point out, however, that these procedures must be applied in relation to context and must evolve as circumstances change.

Bellamy calls his second organizational accomplishment "skeptical standardization." Standardization may serve best when applied to instructional materials, technology, professional development, data systems, and other resources. Standardization is particularly effective in realizing economy and efficiency. As he points out in Chapter Three, an area for the use of standard operating procedures in education is establishing structures and routines that support collective decision-making by teacher teams. Professional learning communities can provide a platform for such

structures and processes to exist. Installing SOPs such as protocols for reviewing student work, monitoring progress of individual students, and collectively responding to the first signs of failure, may provide one of the most promising applications of high-reliability processes.

Design structures and processes for defined autonomy and constrained improvisation

In their study of district-level leadership, Waters and Marzano (2006) discovered a "surprising and perplexing finding: one study in their meta-analysis found that building autonomy was positively correlated to student achievement in the district; but that same study reported that *site-based management* exhibited a neglible or even negative effect on student achievement. The authors resolved these seemingly contradictory findings by coining the term *defined autonomy* (Waters & Marzano, 2006; Marzano & Waters, 2009). The essence of defined autonomy is that "the superintendent provides autonomy to principals to lead their schools, but expects alignment on district goals and use of resources for professional development" (2006, p. 16). One of the associated practices for defined autonomy that surfaced from the research was that superintendents and district staff recognize that a key function is "allowing for and promoting innovation at the school-level within the context of district goals" (p. 16). Defined autonomy actually resides at multiple levels in educational systems, particularly in the relationship of districts to schools and in the balance of school-level goals and procedures with teacher freedom and flexibility in the classroom. Bellamy calls his third organizational accomplishment "constrained improvisation," which he describes as "a collective process that occurs within constraints of organizational values, collaboration, and previously established decision-making routines (see p. 28).

The seemingly paradoxical characteristics of defined autonomy and constrained improvisation lie at the heart of mindful educational organizations as they strive toward higher reliability. This characteristic of HROs could be considered the yin to the yang of standardization. It captures the dual operating modes of centralized procedural control and the necessary organizational flexibility to shift decision making to those closest to the action. In order to attain this flexibility, HROs deliberately sustain diverse perspectives and expertise.

HROs often refer to the on-the-ground improvisation as "work-arounds." While in-the-moment sensitivity to operations is critical to high reliability functioning, work-arounds sometimes can create what is referred to as a *drift away* from standard operating procedures and effective performance. Thus, work-arounds continue to be a concern, even in traditional HROs such as air transportation and chemical safety. In fact, at the most recent International Conference for High Reliability Organizing, how to manage work-arounds was one of the most frequently discussed topics. The concern is the ability of HROs to differentiate between being flexible and "freelancing" too loosely in the moment and to learn from work-arounds—possibly even incorporating new and better practice into standardized procedures.

Almost paradoxically, for schools it may be those structures and processes for determining when to shift from normal operations to improvisation that most needs to be clearly delineated in terms of SOPs. Expanding upon the use of protocols among teacher teams, there could be clear procedures for identifying students at the first indications of failure. Trigger mechanisms for shifting a response from the teacher to a team with a diversity of expertise could be clearly articulated. Indeed, this is exactly what schools did in the HRS project that Stringfield describes in Chapter Two.

Create and maintain safe reporting cultures ("just" cultures)

HROs create conditions in which employees notice and communicate about early-stage problems that threaten reliability (and safety) and establish incentives for reporting, even if the reporting is done by the individual who made the error or mistake. HROs are constantly concerned with establishing and maintaining safe reporting cultures, or what many of them refer to as "just" cultures. An excellent example of guidelines to create such conditions in the health care profession appears in the "Principles of a Fair and Just Culture" from the Dana Farber Institute in its *Patient Safety Rounds Toolkit* (2004). The seven principles outlined in this document are based upon a core value that "in order to have the greatest impact and achieve the highest level of excellence, staff must be able to speak up about problems, errors, conflicts and misunderstandings in an environment where it is the shared goal to identify and discuss problems with curiosity and respect" (p. 1).

It may very well be those in the organization with an out-of-the-box perspective are best suited to identify the weak signals of impending failure that the rest overlook. For Stringfield and his colleagues, it also means honoring the flaw finders. HROs respect the opinions of even those who find fault, despite being an occasional thorn in leadership's side.

"Public warnings and orderly transitions" is Bellamy's fourth HRO accomplishment. Unlike typical HROs,

where problems that threaten reliability are infrequent, Bellamy reminds us that schools face emerging teaching and learning problems all the time. Individual work-arounds and near misses occur constantly, but in too many schools and districts, there is no systematic approach, and lessons learned don't make it to the collective level and become SOPs.

Weick and Sutcliffe (2001) briefly describe the critical yet tentative elements of "credibility" and "trust," both intangible but essential assets. People must feel safe to be able to identify errors in the system, even if they are the ones who commit them. Of course, the organization must follow up on any reports and take some type of action, even if it is simply an acknowledgement. Never should the flaw finders be ostracized or punished.

Lessons from medicine

During the past 10 years, the medical profession has been actively applying lessons from HROs to reduce errors and mistakes that, if left unchecked, can lead to unnecessary patient suffering and death. Weick and Sutcliffe (2007) cite a report on medical errors by the Institute of Medicine that concluded: "[Health care] is very different from a manufacturing process, mostly because of huge variability in patients and circumstances, the need to adapt process quickly, the rapidly changing knowledge base, and the importance of highly trained professionals who must use expert judgment in dynamic settings" (p. 39). Their description similarly applies to educational systems, and as a result, the application of HRO principles, characteristics, and strategies to health care may provide us with guidance.

Several authors, many of whom are physicians and surgeons themselves, have written about ways the medical profession is seeking to improve their process reliability from doctor to doctor and among hospitals. For instance, in *Better: A Surgeon's Notes on Performance*, Atul Gawande (2007) describes how hospitals maintain a continual focus on the prevention of failure by implementing overlapping protocols to decrease the possibility of mistakes. The health care profession has begun to explore the application of high reliability concepts to areas in which failure of the system does not result in immediate death, but instead on the "slow burning" events such as infection rates, heart disease, obesity, and long-term care.

In both education and health care, we face the constant challenges that arise from the complexity of human beings, physically and cognitively, of the human frailties of doctors, nurses, teachers, and administrators, and of the human interactions between provider and recipient. Resar (2006) identifies four themes in health care settings that help to explain at least a portion of the gap in process reliability:

1. Extreme dependence on hard work and personal vigilance.
2. Focus on mediocre benchmark outcomes rather than process.
3. Great tolerance of provider autonomy.
4. Failure to create systems that are specifically designed to reach articulated reliability goals.

Resar concludes that "the resulting variability in the process of delivering care forces the organization in which these autonomous providers work to develop a supporting infrastructure that is at best marginally effective" (p. 1683). Is it any wonder we in education are experiencing many of the same issues? Optimistically, lessons from High Reliability Organizations may provide both professions with routes to improvement.

Lessons learned from failure

Some of the most powerful lessons learned from HROs come from failure. Even though NASA adopted the slogan "failure is not an option," the Challenger and Columbia space shuttle disasters still occurred, resulting in the deaths of 13 talented astronauts, one of whom was to be the first teacher in space. Inquiries into the causes for these incidents found that technical failures and cascading human errors and mistakes were compounded by organizational culture.

More recently, we have seen a similar pattern in the BP oil spill (see Chapter Three for Bellamy's analysis). The natural disaster of the Japanese earthquake and tsunami on March 11, 2011, and the near-meltdown of the Fukushima nuclear power plant, which resulted from multiple failures in prevention and in the design of backup systems, is providing an extraordinary learning opportunity for the nuclear power industry. We can hope that the good news will be that not only do the individual organizations that were directly affected learn by these disasters, but so will entire industries.

As Bellamy concludes, organizations most often fail when goals for performance or achievement supersede goals for avoidance of critical errors. The attention to the bottom line for shareholders surpasses attention to safety. The organizational culture shifts from one that encourages error reporting to one that demands compliance and punishes whistle blowers and those with different perspectives. A recent educational equivalent comes to mind: The investigation into the Atlanta Public Schools cheating scandal. Observations from an *Atlanta Journal-*

Constitution article (Vogell, 2011) include these:

- Across Atlanta Public Schools, staff worked feverishly in secret to transform testing failures into successes.
- Teachers and principals erased and corrected mistakes on students' answer sheets.
- Area superintendents silenced whistle-blowers and rewarded subordinates who met academic goals by any means possible.
- Superintendent Beverly Hall and her top aides ignored, buried, destroyed, or altered complaints about misconduct, claimed ignorance of wrongdoing, and accused naysayers of failing to believe in poor children's ability to learn.

James Reason (2000), another physician, tells us that High Reliability Organizations are not immune to adverse events, but they are able to convert these occasional setbacks into enhanced resilience of the system. Whether or not public education systems in the United States are resilient enough to bounce back from repeated events like this is in question.

Some encouraging news—the latest McKinsey study

*How the World's Most Improved School Systems Keep Getting Better (*Mourshed, Chijoke, & Barber, 2010) examines 20 school systems from around the world, all of which reported significant, sustained, and widespread gains in student performance on international assessment measures, but each at a different stage in its improvement trajectory. The authors of this new report map out a journey of improvement along different stages of the performance spectrum—from poor to fair, fair to good, good to great, and great to excellent.

The report finds that at early stages (i.e., poor to fair, fair to good), the systems dictate "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" (p. 52).

As systems move upward toward good to great, they are characterized by more highly skilled educators. They provide only loose guidelines on teaching and learning processes because peer-led creativity and innovation inside schools becomes the core driver for raising performance. Standard operating procedures are relaxed and the system moves from tighter to looser control. However, at all stages, the systems focus attention on a few non-negotiable key goals.

Once again, this report emphasized the importance of leadership at various levels of the system. To initiate the change toward improvement, leadership transition was necessary, but once the trajectory was established, leadership stability became very important.

A follow-up to the 2007 study of the world's highest performing systems, this report verifies that many of the HRO principles, characteristics, and strategies we have synthesized in this monograph are being implemented in the world's highest performing systems and in those on successful improvement trajectories. We believe that, by understanding the characteristics and adhering to the key principles of high reliability, education can achieve higher levels of performance and lower levels of failure. For leaders and practitioners, the questions become, "Will we achieve reliability at the levels that true HROs operate?" and "Should we commit to anything less?"

Challenging our assumptions about schooling

Jared Diamond, in his book *Collapse: How Societies Choose to Fail or Succeed* (2006), writes: "Perhaps a crux of success or failure as a society is to know which core values to hold on to, and which ones to discard and replace with new values, when times change" (p. 433). At McREL, we have adapted that statement to this: Perhaps the crux of success or failure of American education is for leaders to know which practices to hold on to, which ones to discard, and how to significantly improve execution of effective research-based practices, as times and external demands change.

A bold new initiative

As a result of the feedback from our initial exploratory event in October 2010, McREL is launching a national "best in the world" consortium of leaders from high-performing U.S. schools and districts that, together, will work to reverse the downward slide of U.S. schools by raising both the "floor" and the "ceiling" of student performance. The Network for Innovative Education is an initiative to reduce the achievement gap not only between low- and high-achieving students in the United States but also between the highest performing systems in the United States and the "best in the world."

References

Armario, C. (2010, December 7). *'Wake-up call': U.S. students trail global leaders.* Retrieved from http://www.msnbc.msn.com/id/40544897/ns/us_news-life/t/wake-up-call-us-students-trail-global-leaders

Barber, M. (2010, October 28). *International lessons in education reform.* Presented at the Best in the World Exploratory Gathering, Denver, Colorado.

Dana-Farber Cancer Institute. (2004). Principles of a fair and just culture. *Patient Safety Rounds Toolkit.* Retrieved from http://www.dana-farber.org/uploadedFiles/Library/adult-care/care-quality-and-safety/patientsafety-resources/patient-safety-rounding-toolkit/principles-of-a-fair-and-just-culture.pdf

Diamond, J. (2006). *Collapse: How societies choose to fail or succeed.* London: Penguin Books.

Gawande, A. (2007). B*etter: A surgeon's notes on performance.* New York, NY: Picador.

Goodwin, B. (2010). *Changing the odds for student success: What matters most.* Denver, CO: Mid-continent Research for Education and Learning (McREL).

Hanushek, E. A., Peterson, P. E., & Woessmann, L. (2010, November). *U.S. math performance in global perspective: How well does each state do at producing high achieving students?* Retrieved from http://www.hks.harvard.edu/pepg/PDF/Papers/PEPG10-19_HanushekPetersonWoessmann.pdf

Hoy, W. K. (2003). An analysis of enabling and mindful school structures: Some theoretical research and practical considerations. *Educational Administration Quarterly, 41*(1), 87–108.

Hoy, W. K., Gage, C. Q., & Tarter, C. J. (2006). School mindfulness and faculty trust: Necessary conditions for each other? *Educational Administration Quarterly, 42*(2) 236–255.

Langer, E. J. (1989) Mindfulness. Reading, MA: Addison-Wesley/Addison Wesley Longman.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). C*lassroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., & Waters, T. (2009). *District leadership that works.* Bloomington, IN: Solution Tree.

McKinsey & Company. (2007). *How the world's best performing school systems come out on top.* New York, NY: Author.

McKinsey & Company. (2009, April). *The economic impact of the achievement gap in America's schools.* New York, NY: Author.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better.* New York, NY: McKinsey & Company.

OECD. (2010). PISA 2009 results: *What students know and can do–Student performance in reading, mathematics and science. (Volume 1).* Retrieved from http://www.oecd.org/document/53/0,3746, en_32252351_46584327_46584821_1_1_1_1,00.html

Perrow, C. (1999). *Normal accidents: Living with high risk technologies.* Princeton, NJ: Princeton University Press.

Reason, J. (2000, March 18). Human error: Models and management. B*ritish Medical Journal, 320*, 768–770.

Resar, R. (2006). Practical applications of reliability theory: Making noncatastrophic health care processes reliable. *Health Services Research, 41*(4), 1677–1689.

Roberts, K. H. (1990). Some characteristics of one type of high reliability organization. *Organization Science, 1*(2), 160–176.

Schleicher, A. (2010). *The high cost of low educational performance.* [Video]. Available from http://eux.tv/2010/01/andreas-schleicher-on-the-high-cost-of-low-educational-performance/

Senge, P. (2006). *The fifth discipline: The art and practice of the learning organization* (rev. ed.). New York, NY: Doubleday.

Vogell, H. (2011, July 6) Investigation into APS cheating finds unethical behavior across every level. *Atlanta Journal-Constitution.* Retrieved from http://www.ajc.com/news/investigation-into-aps-cheating-1001375.html?printArticle=y

Waters, J. T., & Marzano, R. J. (2006). *School district leadership that works: The effect of superintendent leadership* on student achievement. Denver, CO: Mid-continent Research for Education and Learning.

Weick, K. E. (2011, April 20*). Putting HRO into practice.* Presented at High Reliability Organizing Conference, Washington, DC. Retrieved from http://www.high-reliability.org/Presentations/Conferences/Washington_DC/Weick_Hro_2011_revapril7_outline_and_speech.pdf

Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected: Assuring high performance in an age of complexity.* San Francisco, CA: Jossey-Bass.

Weick, K. E., & Sutcliffe, K. M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty.* San Francisco, CA: Jossey-Bass.

Jim Eck is a senior director in McREL's Field Services. His expertise is in leadership, standards, assessment, and facilitating change efforts within state and local education agencies. He can be contacted at jeck@mcrel.org or 303.632.5517.

McREL

Mid-continent Research for Education and Learning

4601 DTC Blvd., Ste. 500, Denver, CO 80237-2596
P: 800.781.0156 • F: 303.337.3005
Website www.mcrel.org • E-mail info@mcrel.org