Learning to Improve
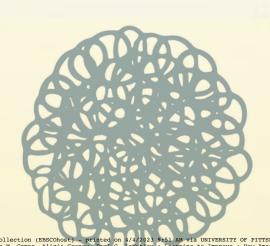
How America's Schools Can Get Better at Getting Better

Anthony S. Bryk

Louis M. Gomez

Alicia Grunow

Paul G. LeMahieu

Preface

IN THE FALL OF 2007, the American Enterprise Institute organized a conference titled the "Supply Side of School Reform" in Washington, DC. The conference created an occasion for two of the authors of this volume, Tony Bryk and Louis Gomez, to draw together a set of ruminations, based on over a decade of collaboration on work to improve student outcomes in the Chicago Public Schools. The "Ruminations" paper offered a sober appraisal.[1] Something was seriously wrong with the ways the field of education sought to connect research to practice improvement. One could see small successes here and there, but the overall processes of improvement were very fragile and much too slow. The forces at work in large urban school systems were competing tsunamis of change. Each hot new idea seemed well intentioned, but often impervious to evidence. The cumulative effect of various reforms, layered one on top of another, was often less than helpful. It seemed clear that if educational reformers continued to do what they had always done, education would continue to get more of the same—great variability in outcomes that often further disadvantages the most disadvantaged in our society.[2]

We questioned how our nation could possibly improve our schools without a transformation in the ways it develops and supports school professionals and the materials, ideas, and evidence with which they work. We noted that when other sectors of society confront mounting challenges, such as those facing education today, leaders naturally turn

to their research and development communities for guidance. Unfortunately, the research and development infrastructure for school improvement is weak and fragmented. The core institutional arrangements among public education, universities, and the commercial sector are failing our schools. Moreover, even when constructive efforts operate, they occur within a political environment that continues to seek quick fixes rather than investing in the sustained work necessary to advance quality outcomes reliably at scale.

Our universities reward faculty members for their scholarly research contributions; the individually authored paper in a refereed journal is the prized accomplishment. This institutional culture and its academic incentives are not conducive to teams working together on the kinds of design, development, and refinement activities that are necessary to solve educational problems.[3] To be sure, valuable knowledge is emerging from schools of education and across the social sciences. This knowledge has salience and *could* improve schooling significantly. It is just not happening at the speed and scope of what is needed and is possible.

Likewise, while individual educators may develop considerable knowledge through their daily work, no mechanisms exist to test, refine, and transform this practitioner expertise into a professional knowledge base.[4] School districts are not especially proactive in developing and improving instructional materials, practices, and programs based on careful design and testing. Systematic evidence-based improvement is just not in the fabric of their work. Instead, educational leaders move quickly to implement untested solutions in responding to ever-changing government and foundation initiatives. Educators are constantly running from behind, seeking to keep up before the next new reform—often disconnected from the last reform—sweeps over them.

Not surprisingly, the commercial sector is also lacking in this context. It responds quickly to changes in policy and funding priorities by bringing forth many new products and services, but individual firms rarely initiate and commit to the kind of sustained R&D necessary to improve schools. There is just too much market uncertainty to warrant such longer-term risk taking.

The Ruminations paper sketched a design for an R&D enterprise that would specifically focus on improving our nation's schools and colleges. In brief, the paper argued that we needed to

- Focus on improving the actual day-to-day work in classrooms, schools, and districts
- Orchestrate diverse expertise among researchers, practitioners, designers, and developers to solve practical problems
- Embrace an approach that focused on multiple quick tests of change and iterative refinement of promising ideas
- Recognize that variation in organizational contexts is a core design and development challenge rather than just some externality to be ignored
- Reframe the overall mission as accelerating how a field learns to improve its core work

All of this seemed very sensible to us, yet very little research was actually organized this way. We wondered whether these ideas could actually work.[5]

And then something unexpected happened.

Inspiration—seeing a complex problem through fresh eyes—arises on occasion in most unexpected ways. In the summer of 2008, we were approached by Jim Kohlmoos, then the executive director for the Knowledge Alliance, a Washington-based association of major applied research organizations in the United States. Some of its members had read Ruminations and were intrigued. The Knowledge Alliance invited us to present at a conference it was organizing for that December at Snowbird, Utah. It was there that we first met Donald Berwick, then-president and cofounder of the Institute for Healthcare Improvement (IHI). As Don talked, we realized that this organization, which had been around for nearly twenty years, was actually living the ideas and doing the kind of work in the health-care sector that we had been speculating about for education. Don, who was trained as a pediatrician, subsequently invited us to visit IHI to see all of this firsthand. Little did we know at the time that he had just opened a door for us to an extraordinary high-performing organization.

At the conclusion of our first visit to IHI, Don shared a copy of *The Best Practice* by Charles Kenney.[6] The book offers a narrative account of the emergence of improvement research in health care. Over two decades ago, a few extraordinary leaders in health care recognized the challenges confronting their field. These leaders borrowed the discipline and culture of quality improvement, pioneered by Japanese industry, and then applied it to the complex enterprise that is modern health care.[7] Under their influence, medical institutions have reduced harm, saved lives, and enhanced patients' experiences while also controlling costs.[8] Their efforts to promote a system in which "continuously improving is what we do here" became our inspiration.[9] At several points in Kenney's text, one could easily substitute the words *teachers* and *students* for *doctors* and *patients* and believe that this was actually a discourse about education. We were convinced that if those in health care could learn useful lessons from studying industrial quality improvement, there was a good chance that educators could learn something in turn from them.

So we became serious students of IHI—how it organized its work and why and how it had come to do so as it did.[10] IHI generously opened its doors to us. We have participated in several of its professional training programs. Its staff and associates have been our mentors as we sought to adapt to education many of the principles, tools, and routines of improvement science that have been so successfully deployed in health care.

Along the way, I (Tony) had the privilege of being invited to serve on an Institute of Medicine committee. Through these discussions, I came to understand better why leaders in this field were seeking to promote a more dynamic and vigorous learning-to-improve health-care system. The medical field attracts bright, well-trained professionals and compensates them well. It draws from a vast reservoir of basic and applied research and cutting-edge technology applications. But even with the best of people, tools, and technical knowledge, tremendous variability exists in health-care outcomes. Many hospitals aren't as good as they could be because they focus too little attention on how to bring their talent and technology together effectively day-in and day-out on behalf of patients. Their systems do not support their many professionals effectively, and they do

not sufficiently address the complex and often very trying conditions under which they work.[11]

As an organizational sociologist who has spent many years seeking to improve urban schools and districts, the implications seemed obvious. Education too has a serious learning-to-improve problem. Even as we await better knowledge, tools, and other resources, we must accomplish more with what we already know and the resources we already have. We believe this is achievable in education, as is now occurring in select health-care institutions, when leaders seriously commit to systematic quality improvement.

ANALOGICAL SCAVENGERS

Our deep dive into health-care quality improvement shaped a practical orientation that continues to guide our efforts. We authors of this volume have come to think of ourselves as a new breed of birds, the "analogical scavenger." We are constantly looking to other fields that share a concern about improving practices and that have made some significant progress. We study these deeply, reflecting on commonalities with the education field, while also scrutinizing critical differences. How do these ideas fit? Where are adaptations needed, and how do we discern which are most appropriate?

Learning to Improve draws on some of the best ideas emerging in education itself in the form of communities of practice, teacher action research, lesson study, and the scholarship of teaching and learning.[12] We also acknowledge close colleagueship with a set of practices labeled developmental evaluation.[13] Likewise, we are indebted to a growing scholarship around user-centered and design-based implementation.[14] We have also drawn liberally on the practical strategies developed in commercial design firms, such as IDEO, and in related academic centers such as the Hasso Plattner Institute of Design at Stanford. In each instance, we have scavenged for good ideas and practices that have worked in other contexts but may not yet have received their due attention in education.

Consequently, we owe debts to many others for the basic principles and practical tools detailed in this book. Our main contribution in this regard

is as an integrative agent, drawing together the best of what we have found and melding it into a coherent system of principles and methods for accelerating learning to improve. Ultimately, this led us to the concept of *networked improvement communities* (NICs). This new form for educational R&D joins together the discipline of improvement science with the dynamism and creative power of networks organized to solve common problems. NICs are about helping America's schools get better at getting better.

LEARNING-BY-DOING

In tandem with our scavenging activity, we also set out to test and refine these ideas, methods, and tools by seeking to use them to address important improvement problems in colleges and schools. One of our goals in the book is to share our early experiences at the Carnegie Foundation for the Advancement of Teaching in attempting to make two NICs come alive.

In 2010 we initiated a Community Colleges Pathways NIC focused on the extraordinarily high failure rates in developmental mathematics courses in community colleges. Unable to acquire necessary college credits, these students cannot transfer to a four-year institution nor qualify for entry into specialized technical and occupational training programs. These courses literally function as a gatekeeper to opportunity for hundreds of thousands of students every year. By July 2011 some twenty-seven colleges had joined the Pathways NIC to attack this problem. Where historically only 5 percent of the students assigned to developmental math classes received college math credit in one year, 50 percent of students participating in the Statway® and Quantway® programs developed by the network now achieve college mathematics credit in a year. Measured improvements have emerged in virtually every participating college and for every subgroup of students. These success rates have been sustained over three years even as the NIC has expanded to more than fifty colleges and to many more faculty. The network actively continues to learn from its data in efforts to further improve its students' outcomes.

In the fall of 2011, the Carnegie Foundation launched a second and smaller NIC focused on the weak systems that bring new teachers into

public schools, fail to support them in learning to teach well, and consequently under-educate the children in their classrooms. These systems function as a revolving door where many new teachers leave the profession and the process just recycles with a new group over and over. District teams in Baltimore, Maryland; Austin, Texas; and the New Visions for Public Schools Network in New York City joined together to attack this problem. While each district initially pursued some different change ideas, the network eventually focused on two processes that strongly affect new teachers' lives: the quality of feedback that new teachers receive for improving their teaching and the degree of support that new teachers sense in their relationship with their primary supervisor, the school principal. New feedback and support processes were developed, tested, refined, and adapted to work reliably across different kinds of schools. A measurement infrastructure is now in place that allows local improvement teams to track process measures about the frequency and quality of these interactions and how this in turn connects to teacher reports about job satisfaction and burnout and to key longer-term outcomes—measures of teaching effectiveness and decisions about possibly leaving teaching. This network has directly engaged teachers, principals, and other school-based educators in improvement research. In our view, the improvement paradigm that we introduce in this book has passed a key practitioner test. Increasing numbers of school leaders in these districts now want to learn how to use improvement methods to address a broader array of local problems.

These two NICs began somewhat differently and as a result provide complementary and contrasting experiences. The Community College Pathways NIC formed quickly as a network around the design, development, and scaling of new courses of instruction in developmental mathematics. It offers key insights about the early stages of network formation. The second NIC, which we call Building a Teaching Effectiveness Network (BTEN), began by immersing district leaders and school principals in rapid, small tests of change seeking to improve the quality of feedback and support that new teachers receive. It offers key insights about the mechanics of introducing such disciplined inquiries into day-to-day school affairs. Interwoven throughout this book are brief illustrations of select

tools and processes now being used by these two NICs. We also detail some issues that have surfaced as educators begin to take up active roles as improvement researchers and as academic researchers attempt to interact in very different ways with their clinical colleagues.

Learning to Improve seeks to make these ideas and this emergent know-how accessible to a larger audience. This book is merely a starting point. It is an introduction to a new way of tackling problems in education and an invitation to join in the journey of building the capacity for quality improvement in our nation's schools and colleges.

Introduction

A Better Way

The history of American education includes a graveyard
of good ideas condemned by pressure for fast results.

—JAMES HIEBERT, RON GALLIMORE, AND JIM STIGLER[1]

BY THE LATE 1990s, many policy advocates and educational reformers
had concluded that U.S. high schools were too big and too impersonal.
In attempting to be everything to everyone, comprehensive high schools
failed many students, especially those from disadvantaged backgrounds.
Anonymity bred apathy and alienation among students and teachers alike.
Far too many students dropped out, and many of those who did persist to
graduation were ill prepared for work or higher education.

A reform movement emerged. One of its most articulate advocates was
Tom Vander Ark, the senior program officer for education at the Bill &
Melinda Gates Foundation. Testifying before a congressional committee
in 2001, he pressed for smaller, more personalized school settings, places
where faculty and students would really know each other, leading to a
strong sense of shared commitment to each other and to academic suc-
cess.[2] Vander Ark described how visionary leaders at Central Park East
and Urban Academy in New York City and other small schools had cre-
ated impressive examples of truly engaging learning environments.[3] These

compelling stories were buttressed by a growing body of research showing that high schools with fewer than 400 students produced significant benefits, including higher student attendance and graduation rates; improved school climate and safety; greater parent and community involvement; and higher staff satisfaction. Moreover, research found that smaller schools benefitted economically and socially disadvantaged students the most.[4] Vander Ark urged federal leaders to promote the redesign of the American high school.

The Gates Foundation pursued that goal aggressively over the next several years. It spent some $2 billion promoting the dissolution of large high schools and the creation of some 2,600 smaller ones in forty-five states and the District of Columbia.[5] New York City alone created more than two hundred such schools. The scope and pace of change was breathtaking. Embedded here is a remarkable story of how a single philanthropic institution leveraged its resources to transform the institutional landscape of American public education in a few short years.

Yet all was not well. While the number of small schools grew at an astounding rate, so did the problems. Only a small number of urban educators had ever worked in small high schools, let alone started them. Even fewer had taken on the more difficult task of transforming large dysfunctional high schools into multiple smaller units that could coexist under a single roof. Few educational leaders knew how to maneuver effectively through the swamp of day-to-day problem solving required to bring this new institutional form to life.

Equally problematic were the on-the-ground politics of change. Many school staff were resistant to the small school idea. They saw it as just one more instance of change being imposed from outside by reformers who knew little about the conditions that teachers and principals actually confronted. The human and social resources necessary to engage meaningful change of any sort were also lacking in many of the Gates-funded schools. Teachers reported extremely low levels of satisfaction in their jobs, little inclination to try new things, and little trust in each other.[6] Breaking large comprehensive high schools into smaller units also challenged the long-standing authority of department heads and other leaders in tradi-

tional high schools. Not surprisingly, there was much wrangling over de-tails that proved time-consuming and divisive.[7]

Most significant, many of the new, smaller high schools that emerged from the Gates initiative developed in ways very different from places like Central Park East and Urban Academy that had been held out as models. Those successful small schools had started with a select group of faculty; built up gradually by adding a grade each year; and learned from trial and error, changing and improving over time.[8] In essence, their leaders brought what design professionals refer to as a design and development orientation to the task of new school creation.[9]

In contrast, under the Gates initiative, districts often compelled staff in large existing high schools to redesign themselves into clusters of small schools to be housed in the same facility. School faculties who were already working under exceedingly difficult conditions and who had no prior experience in school redesign were charged with essentially reinventing themselves, often in less than a year. There was little opportunity to start small, fail, learn, and iterate toward success. Rapid large-scale change was being called for under conditions least likely to produce success.[10]

So a good idea found itself embedded in a bad development strategy with weak collective will, limited capacity to execute, and an unrealis-tic timetable. Educators across the country were confronting common challenges and trying to solve complex problems. Many learned valuable lessons that could have propelled the small schools initiative toward suc-cess. Unfortunately, most were working on this effort largely on their own. There was no organized system to capture their learning, refine it, and transform it into a collective force accelerating wider-scale improvements. That many small schools struggled and some failed, especially early on, is not surprising.[11] This reform was ambitious and quite novel, and much of the practical knowledge required to make it work was unknown at the outset. That failures were likely to occur under these circumstances is not a fatal flaw; but that the field failed to learn quickly from these failures was, and is.

Although some notable exceptions did subsequently emerge, such as the New Century High Schools in New York City, by late fall of 2008 the

Gates Foundation came to acknowledge that its efforts to rapidly develop small high schools were not the panacea that it had hoped.[12] Another promising reform idea had failed to deliver.

With this effort in the rearview mirror, the foundation moved on to another big transformative idea to fix education.

> [The Gates Foundation] has shifted its considerable weight behind an emerging consensus—shared by U.S. Education Secretary and Gates ally Arne Duncan—that quality of teaching affects student performance and that increasing achievement is as simple as removing bad teachers, identifying good ones, and rewarding them with more money.[13]

A pressing new concern had come into view. Media accounts, such as the "Rubber Room" article in the *New Yorker*, focused attention on teachers whom districts appeared unable to fire.[14] In a full-length motion picture, *Waiting for Superman*, Americans were told that test scores were low because there are so many bad teachers.

Policy leaders quickly jumped on a new reform bandwagon: rigorous teacher evaluations tied to financial incentives and employment decisions. A complex statistical procedure called value-added analysis, which previously had been employed by only a very select group of microeconomists and statisticians, quickly became the centerpiece for these new evaluation systems.

As with small high schools, this reform idea, too, was anchored in a growing body of academic research. In this instance, numerous studies had documented wide variability in student learning among different classrooms. Less clear was how to interpret these findings and to discern their practical implications.[15] Regardless, policy makers quickly came to believe that the new teacher evaluation protocols coupled with new data systems would now be able to tell district leaders who to fire and who to reward.

But here, too, reformers got ahead of themselves. States and districts, supported by resources from both the Gates Foundation and the U.S. Department of Education, moved quickly to implement teacher evaluation procedures even though very significant technical and logistical issues remained unsolved.[16] As with high school redesign, reformers had rapidly

launched a major change strategy in public education with a lack of requisite knowledge, skill and organizational capacity, and not surprisingly with considerable resistance from teachers.

THE CHRONIC FAILURE OF PROMISING REFORM IDEAS

While teacher evaluation and high school redesign are high-profile cases, they are not anomalies. Over and over, change efforts spread rapidly across the education landscape, despite an absence of knowledge as to how (or even whether it is possible) to effect improvements envisioned by reform advocates.

When reformers focused attention on the generally poor quality of professional development efforts to help teachers improve their classroom teaching, a new organizational role—the instructional coach—was introduced into schools.[17] However, what coaches actually needed to know and be able to do, and the requisite work conditions necessary in schools for them to do their jobs successfully, was left largely unspecified. When reformers recognized the importance of principal leadership, principals were asked to take on an expanded role as instructional leaders even though demands on their time were already excessive.[18] When policy makers were unsatisfied with the rate of school improvement, high-stakes accountability schemes were introduced, but unintended consequences abounded, some of them hurting the very students the reforms were designed to help.[19] Reaching back a bit further, when corporate downsizing was the rage some years ago, school districts moved quickly to embrace the concept of site-based management. But, the roles and responsibilities of newly empowered school-based decision makers were often left unclear and the resources needed for carrying out school-based decisions often lacking.[20]

In each instance there was a real problem to solve, and in most cases there was at least a germ of a good reform idea. Educators, however, typically did not know how to execute on the ideas. Districts and states lacked the individual expertise and organizational capacity to support these changes at scale, and policy makers regularly ignored arguably the most

important instrument for any of this to work: engaging the minds and hearts of our nation's teachers and principals on behalf of the reforms.

Educational leaders continue down this path today, believing that they must disrupt the educational system substantially and quickly. They expend great energy rolling out large-scale changes across whole districts and, sometimes, whole states.[21] These initiatives make extraordinary demands on leaders' time, as they seek to advance broad changes while also working hard to sustain political support in the face of inevitable implementation problems. Teachers, principals, students, and parents are taxed too. Directives change, guidance is absent, and key provisions sometimes don't work at all. When the data are crunched, the same disappointing conclusions emerge. The press to quickly push good ideas into large-scale use rarely delivers promised outcomes. Results are typically modest and vary from school to school. In some locales a reform might work; in many places it does not; and in some instances it might even do harm.

At base is a common story of *going fast and learning slow*. We consistently fail to appreciate what it actually takes to make some promising idea work reliably in practice. We become disappointed when dramatic positive results do not readily emerge, and then we just move on to the next new reform idea. This should trouble all of us. If we continue to seek improvement in the ways we have always done, we are likely to continue to get what we have always gotten.

IMPROVEMENT SCIENCE AND NETWORKED COMMUNITIES

Each reform case, mentioned previously, sought to implement fast and wide and then fix problems later. This strategy has failed again and again. Comforting in a somewhat perverse way, such results are not peculiar to education. They are also seen in other sectors that have sought to improve their productivity in similar ways.[22] Experiences across many different fields now caution humility about how much must be learned in order to transform successfully a change idea into new human capabilities, into day-to-day practices that work reliably, and into the redesign

of organizational arrangements necessary to support all of this. Achieving successful change in complex work systems means recognizing that one cannot predict ahead of time all of the details that need to be worked through nor the unintended negative consequences that might also ensue. This is just an operational fact of life about the nature of complex organizations. And as we detail in chapter 4, contemporary educational institutions are indeed complex.

Understanding this dynamic presses a fundamental shift in how we think and act, a shift toward *learning fast to implement well*.[23] It also calls out for very different organizational arrangements to accomplish this end. As we elaborate across the pages of this volume, the concept of *networked improvement communities* (NICs) offers an attractive alternative. A NIC unites the conceptual and analytic discipline of improvement science with the power of networked communities to innovate and learn together. In embracing improvement science, educators are able to draw upon a well-established set of tools and deep practical experiences. Many different kinds of institutions have learned faster and better by using these methods. In carrying out this activity through networked communities organized to solve a shared problem, it is possible to accelerate improvements even further and to engage actively many different individuals and institutions in the process. Moreover, this strategy is broadly useful whether the target for improvement is the classroom, school, faculty network, or school-community partnership; or a whole college, school district, or state education system.

Improvement Science Disciplines Inquiries

Over the past half-century, notable successes for improvement science have occurred first in industry and then more recently in social sectors such as health care.[24] The lesson those experiences teach is that problems ranging from defective products to hospital-induced infections do not stem primarily from an absence of basic research or inferior workforces. Rather, they result from the ways that work systems are designed and thereby shape how individuals carry out their responsibilities.

Improvement science addresses this reality by focusing on the specific tasks people do; the processes and tools they use; and how prevailing

policies, organizational structures, and norms affect this. Applying improvement science to education would direct greater attention to how better to design and fit together the many elements that shape the way schools work. The latter is key to making our educational institutions more effective, efficient, and personally engaging.

Analytically, improvement research entails getting down into the micro details as to how any proposed set of changes is actually supposed to improve outcomes. Unfortunately, such careful on-the-ground systems thinking rarely characterizes most educational reforms. Typically, a reform's logic of action is vague and almost always underspecified. When such reforms are scrutinized closely, zones of wishful thinking—gaps in understanding, questionable assumptions about causes and effects, and tacit beliefs of the form "and then something good will happen"—regularly abound.[25]

In response, improvement science deploys rapid tests of change to guide the development, revision, and continued fine-tuning of new tools, processes, work roles, and relationships. The approach is explicitly designed to accelerate learning-by-doing. As iterative cycles of change proceed, previously invisible problems often emerge, and improvement activities may need to tack off in some new directions. The objective here is quite different from the traditional pilot program that seeks to offer a proof of concept. Improvement research, in contrast, is a focused learning journey. The overall goal is to develop the necessary *know-how* for a reform idea ultimately to spread faster and more effectively.

This strategy draws on a natural human instinct.[26] Much practical learning occurs every day in schools. Individual teachers learn when they introduce a new practice in their classroom and then carefully evaluate the resulting student work. Likewise, individual schools learn as staff examine data together on the effectiveness of current practices and test improvement ideas against evidence of changes in students' work. Organizations across numerous fields have become much more productive by acknowledging this natural inclination and by building on it in deliberate and systematic ways.

A commitment to empirical evidence anchors this learning orientation. Participants are constantly asking three *core improvement questions*: "What is the specific problem I am now trying to solve? What change might I introduce and why? And, how will I know whether the change is actually an improvement?" Change ideas are tested and refined based on evidence from what actually happened, both intended and otherwise. When a change fails to produce expected results or creates unintended consequences, it forces deeper thinking about what meaningful improvements will actually entail. As subsequent cycles of redesign and testing unfold, a better understanding evolves of the actual problem or problems that need to be solved and more workable interventions begin to emerge.

Improvement science also promotes a different angle of view on efforts to spread effective change. The latter is no longer assumed to be simply a matter of implementing mechanically some processes designed by others. To be sure, well-designed tools and processes matter. But even when such resources are at hand, achieving quality outcomes at scale demands sustained attention to solving a diverse array of local issues. So these too become grist for improvement research.

At its most basic and human level, improvement science is not some set of specialized studies carried out exclusively by external researchers. Learning to improve demands the active, full engagement of educators. This provision challenges prevailing arrangements in which researchers study schooling, design interventions, and analyze policies; and then teachers, principals, and education leaders are cast as *users* of this research in their work.[27] Improvement science, in contrast, brings educators into regular interaction with a broad array of academic and technical experts. Participants in an improvement network form as a colleagueship of expertise—academic, technical, and clinical—deliberately assembled to address specific problems. All involved are now *improvers* seeking to generate strong evidence about how to achieve better outcomes more reliably.

We use in this book a number of related terms that are commonly found in writings about improvement in other fields. Formally, we define them here, even though on occasion we may be a bit more casual in their

use. *Improvement science* is a methodology that disciplines inquiries to improve practice. Undergirding it is a distinctive epistemology about what we seek to learn and how we may come to understand it well.[28] Particular acts of inquiry are *improvement research* projects. These projects aim for *quality improvement.*[29] In the context of education, this refers to the capacity of an organization to produce valued outcomes reliably for different subgroups of students, being educated by different teachers and in varied organizational contexts. Since improvement research is an iterative process often extending over considerable periods of time, it is also referred to as *continuous improvement.* Each term offers a somewhat different angle of view on the overall enterprise, and hopefully all will become clear as we illustrate their application in the chapters ahead.

We also note that in embracing the phrase *improvement science* we aim to distinguish sharply from the language commonly used in educational circles today about "research for policy and practice."[30] Generating better guidance for educational policy and improving work practices to make schooling more effective are very different activities. Each makes different information demands and entails different processes through which such information is developed and its utility assessed. Yoking these two needs together may afford some rhetorical convenience in arguing for public funding, but it also does disservice to building the kind of knowledge that is truly useful for improving schools. We firmly believe that a robust methodology—a highly practical form of rigorous inquiry—exists and can be matched to this most important societal need. In this monograph, we detail its guiding principles, some of its key tools and inquiry processes, and its social organization as a scientific community.

Networks Accelerate Learning

As noted previously, teachers, principals, and educational leaders regularly experiment with new approaches seeking to improve outcomes for their students. Given the size and scope of American education, this experimentation occurs on a grand scale. From this angle of view, the lack of wide-scale improvement in educational practice appears anomalous.

Individual educators and institutions are learning much every day, yet as a field we fail to organize, refine, and build on these lessons.

In other contexts, networks are now forming to attack and quickly solve problems that had once been thought difficult and even intractable.[31] For instance, when a mysterious and deadly new virus emerged, subsequently called SARS, a network of labs working separately, but in regular communication with the World Health Organization (WHO), quickly identified the virus and opened the door for new diagnostic testing and future vaccine development. Multiple investigators, each taking a somewhat different tack on the problem, shared day by day what they were learning. The WHO hub facilitated these exchanges and catalyzed conversations about emerging implications and new hypotheses to examine. What under other circumstances might have taken months or years to discover occurred in just a matter of weeks.[32] Ambitious efforts of this sort are emerging all across the biological and physical sciences. Large networks, engaging diverse participants often including nonscientists, are now mapping DNA, carrying out inquiries into the structure of the universe, and resolving complex mathematical theorems.[33]

These science networks represent a new organizational form, deliberately designed to enable effective collective action on solving complex problems and for developing complex products.[34] By breaking up thorny problems and possible solutions into smaller discrete parts, it becomes possible for many more individuals and organizations to offer meaningful contributions.[35] This strategy capitalizes on the fact that key insights often emerge in unusual places. By working through organized networks, the likelihood increases that these ideas may surface, be systematically examined, and if promising, moved rapidly into testing and refinement. In addition, promising practices emerging in the network are likely to diffuse more rapidly and are further tested and refined as others take them up. Accumulating the practical knowledge generated from these multiple tests of change is essential to making a reform work reliably as diverse individuals engage with it across varied contexts. It provides the strongest assurance that as a change scales, what is implemented is actually an improvement.

A *network hub* plays a key role in structuring and supporting this distributed activity.[36] The hub is responsible for detailing the problem to be solved and for developing and maintaining the coherence of the evolving framework that guides efforts among many different participants. It establishes the processes and norms governing how individuals and groups work together and the evidentiary standards for warranting claims. The hub also provides technical resources and supports the open communication mechanisms necessary to accelerate learning networkwide.

Imagine if these vibrant networks were used to solve practical problems in education. They could harness the field's many efforts at improvement and transform them into collective knowledge building. Fortunately, educators have already taken some steps in this direction. Since the early 2000s interest has grown in activity labeled variously as professional learning communities, communities of practice, and faculty inquiry groups.[37] Typically, these communities coalesce around a common problem. As individuals share their experiences, they stimulate insights among others. Networked science appreciates the value of this social learning and seeks to take it a step further by bringing scientific discipline to bear on how plausible change ideas are detailed, tested, and further refined against evidence. Networked science aims to exploit how the social intelligence of a group can accelerate not just individual learning, but a whole profession's capacity to learn and improve.

THE IMPROVEMENT PRINCIPLES

We elaborate over the course of the next six chapters on the basic ideas sketched out quickly in the preceding sections. Each chapter focuses on a distinct principle. Taken together as a set, these principles represent the foundational elements for improvement science carried out in networked communities. We introduce the principles briefly here.

Make the Work Problem-Specific and User-Centered

Anchoring all activity in a NIC is a specific problem to be solved. Detailing this problem statement often proves harder than it sounds. Initial

versions are typically quite general. For example, early conversations with college leaders, which eventually led to the initiation of the Community College Pathways NIC mentioned in the preface, focused on improving student success in program completion. Framing problem statements in this fashion directs attention toward valued social goals—in this instance the important role that community colleges can and should play in creating opportunities for a better job and better life. Such problem framing affords powerful rhetoric for mobilizing political action but provides little guidance as to what actually needs to happen to improve how these organizations function. It indicts many different problems operating across many different departments and jurisdictions both within and outside the colleges. Improvement research, in contrast, demands focus—"What specific problem or small set of problems are we trying to solve?"

Equally important is examining the problem from the point of view of the user—the person who is experiencing it firsthand. In the community college case, a network initiation team scrutinized the success problem from the point of view of the students. What is it that they experience from the time they arrive at the community college doorstep through to successfully completing a program of study? The team considered a myriad of factors that shaped these trajectories and then zeroed in on arguably the single biggest impediment to student success: the high failure rates in developmental mathematics courses. This became the specific problem to be addressed by the NIC.

Focus on Variation in Performance

Variability in performance is the natural state of affairs in complex organizations. Reducing harmful variation and improving overall quality form the prime targets for improvement efforts. In the context of schooling, this means more consistently producing positive outcomes for diverse students being educated by different teachers and in varied contexts.

Adopting this orientation responds to a common educational finding that change ideas work in some places but not others. It directs attention away from simplistic thinking about solutions in terms of "What works?" toward a more realistic appraisal of "What works, for whom, and under

what set of conditions?" As detailed in chapter 2, a NIC seeks to identify and target for change those key processes where major differences in outcomes take root. Improving how this work is carried out can change the overall distribution of outcomes that ensue.

See the System That Produces the Current Outcomes

Our third principle directs attention to the question, "Why do we continue to get the undesirable outcomes observed?" In developing its improvement agenda, a NIC examines how work is actually carried out in classrooms, schools, and colleges and how larger institutional forces shape this. Results from prior research assist in these efforts (i.e., "What is it that we already know about these factors?"). Likewise, educational practitioners are actively engaged throughout the processes of conceptualizing the problem, examining possible change ideas, and learning about improvement as these changes are tried out.

Adopting a systems perspective makes visible many of the hidden complexities actually operating in an organization that might be important targets for change. It generates an interrelated set of hypotheses that form a working theory of improvement for a NIC. This in turn guides the change efforts to be tested, organizes results from this activity as an accumulating body of evidence, and creates an evolving framework for collective action across an improvement community. Chapter 3 introduces a set of processes and tools to assist in this regard.

We Cannot Improve at Scale What We Cannot Measure

Chapter 4 focuses on the centrality of measurement for improvement. Operationalizing this principle directs us to identify the specific measurable aims that the NIC seeks to accomplish. Regular reporting on these outcomes disciplines the work of the community and holds it internally accountable. Absent continuous feedback of such data, one can easily maintain a belief in the efficacy of one's actions even when the warrant for this remains uncertain or nonexistent. Psychologically, leading improvement requires living on the boundary of belief (about

the importance of what one is trying to accomplish) and doubt (as to whether real progress is occurring). Evidence is essential to operating productively on this boundary.

Informing improvement, however, requires more than just measuring targeted outcomes. This is a direct consequence of adopting a systems view about change. First, it is rare to find an educational intervention that consists of a single action that has direct and immediate effect on some targeted aim. Rather, interventions are typically of the form of a "causal cascade": we need to improve "a" in order to achieve "b," which in turn is essential to accomplishing "c." For example, instructional coaching reforms seek to strengthen the relevance and quality of professional development afforded teachers in order to transform their classroom practices in order to improve students' learning. In many instances, educational interventions actually involve multiple interacting causal cascades of this sort. For example, referencing instructional coaching again, these reforms also have to attend to (a) the selection, (b) the initial training, and (c) the continuing professional development of coaches in order to enhance their expertise in working with teachers in order to increase teachers' expertise in carrying out instruction that results in better student learning. Improvement requires attending to each of the component processes that combine together to determine how well the overall system functions.

Adding a second layer of complication, the ultimate aim for a NIC will often be somewhat removed temporally from the activities that are the immediate targets for improvement. For example, in the Building a Teaching Effectiveness Network (BTEN) mentioned in the preface, the ultimate goal was to improve the retention of effective new teachers. These outcomes can take several years to emerge. The NIC's working theory about how to achieve these outcomes, in contrast, targeted a set of processes that begin within the first weeks and months of employment.

Consequentially, improvement research requires gathering data about the specific processes targeted for change, intermediate outcomes directly linked to these processes, and other key markers on the pathway toward achieving the network's ultimate aims. Unfortunately, such data are not

routinely collected. This opens up a whole new demand for measurement in education, but with a very different purpose—helping educators improve what they do. It also raises significant new logistical issues—how can measurement of this sort be easily woven into the day-to-day work of students and educators rather than added as still one more demand on top of what is an already overburdened work system? Chapter 4 includes an example of such practical measurement and introduces a set of guiding principles and tools to assist in building them.

Use Disciplined Inquiry to Drive Improvement

The methods used in improvement research have been tuned to focus on learning quickly and cheaply. In education, this means minimizing intrusions into ongoing schooling activities (since we expect failures to occur, but we just don't know exactly where), while also generating empirical guidance as to what to try next. Improvement typically entails a sequence of inquiries, where the results from each test of change offer guidance for the next test. Formally, each test is akin to a small experiment; the overall arch of activity is an improvement investigation.

The methodology also offers an explicit process for learning how to scale improvements. Inevitably, as new programs, tools, processes, and roles move out into new contexts, they will need to be changed some in order to be integrated into these contexts. Tackling this problem of *adaptive integration* is a standard aspect of improvement research. As practices that have worked in one or few places move out to more diverse contexts, new improvement cycles are launched. The focus now is on "What will it take to make the intervention work under these new conditions?" This is how improvement research iterates toward quality outcomes reliably at scale.

Interestingly, significant by-products result from this approach to spreading improvement. As this activity proceeds, it is also building organizational resources for broader-based change. The educators involved in the early stages of improvement research become a key human resource in subsequent efforts to spread these changes. They have developed know-

how—that is, how to make some set of changes actually work—and can now teach and mentor others along this same path. In addition, because these same individuals have personally experienced success, they are now poised to become evangelizing leaders, building will with colleagues for wider-spread implementation.

In sum, improvement research consists of a highly integrated set of methods for developing the necessary technical knowledge to transform good ideas into practices that work, building human capabilities necessary for this learning to spread, and directly addressing a major challenge in every improvement effort—building will for change. Addressing each of these is essential to scaling change faster and more effectively.

Accelerate Learning Through Networked Communities

In chapter 6, we explore how a NIC, organized as a scientific community, accelerates broad-based improvements. We detail the organizational resources, formal agreements, and normative understandings shared among participants that are necessary for such collective actions to occur.[38] Drawing on experiences from the two Carnegie-initiated networks and those of others, we describe the relationships that operate among network participants and how practical knowledge develops through their efforts and is taken up productively by others.

Membership in a NIC means placing priority on solving a problem together, rather than pursuing a theoretical predilection, methodological orientation, or personal belief. The latter are all resources, but advancing them is not the primary goal. At the most fundamental level, NIC participation challenges the long-standing norm of autonomy in practice that educators have traditionally prized. It means recognizing that today's problems cannot be solved through isolated individual actions. Each participant holds expertise that is valuable in solving a given problem, but each also recognizes that he or she must join together with others to solve it. Consequently, the life of a NIC entails a profound normative shift. It vitalizes a core belief that we can accomplish more together than even the best of us can accomplish alone.

WHY ALL OF THIS MATTERS NOW

Our nation's schools are, and have been for decades, in a constant state of reform.[39] By many accounts they are actually getting better.[40] Unfortunately, our aspirations for schools are accelerating at a faster rate. Consequently, a growing chasm exists between these noble aspirations and what schools can actually achieve.

Today we ask more of our public schools than ever before. The No Child Left Behind Act of 2001 compelled attention to the learning of all of our nation's students, not just some. The new Common Core State Standards substantially raise the bar as to what this learning entails. So as a first priority, we want our schools to become more effective in advancing deeper learning for all students. Second, we live in a time in which tremendous pressures exist on the public purse. For several decades, education commanded increasing public resources. Now the expectation is that our schools should not only get better results but also do so more efficiently. Third, far too many students remain disengaged, walking out the doors of our high schools and colleges and never completing their education. Turnover among teachers, principals, and superintendents is also unacceptably high, and morale is at an all-time low.[41]

A significant advance on any one of these three aims—greater academic effectiveness, cost efficiency, and human engagement—would be a major accomplishment. Simultaneously improving on all three counts would be extraordinary. Yet this is precisely what our educational institutions must now do.

We return to the example with which we began—the movement toward small high schools. The alienating quality of large urban high schools motivated this reform. Small schools promised, and often delivered, much more engaging environments for both students and teachers. But many advocates for this reform largely ignored questions about how instruction would actually improve in small schools. And then suddenly, when they were confronted with an economic downturn and increasing fiscal pressures, policy shifted toward consolidating smaller schools. It was now argued that larger schools provided access to a wider array of courses and

services and were more cost efficient. But what about the implications for student and faculty engagement? And how would this change actually enable more effective instruction?

Attaining the *Triple Aims of Educational Improvement*—improved effectiveness, greater efficiency, and enhanced engagement—seems inconceivable to us so long as we continue to pursue reforms as we typically do.[42] Success will elude educational leaders and each new cadre of educational reformers unless they and their institutions are equipped with better ways of understanding the practical problems needing address and with more systematic approaches toward their improvement.

This book is about setting out a new path—a more dependable way for educators to improve their schools. Our goal is to unleash the dynamism of networked improvement communities. They are our best hope for crossing the growing quality chasm in education.