

Genuine Progress, Greater Challenges:

A DECADE OF TEACHER EFFECTIVENESS REFORMS

Andrew J. Rotherham

Ashley LiBetti Mitchel



BELLWETHER
EDUCATION PARTNERS

IDEAS | PEOPLE | RESULTS

MAY 2014

© 2014 Bellwether Education Partners

This report carries a Creative Commons license, which permits noncommercial re-use of content when proper attribution is provided. This means you are free to copy, display and distribute this work, or include content from this report in derivative works, under the following conditions:

Attribution. You must clearly attribute the work to Bellwether Education Partners, and provide a link back to the publication at <http://bellwethereducation.org/>.

Noncommercial. You may not use this work for commercial purposes without explicit prior permission from Bellwether Education Partners.

Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under a license identical to this one.

For the full legal code of this Creative Commons license, please visit www.creativecommons.org. If you have any questions about citing or reusing Bellwether Education Partners content, please contact us.

TABLE OF CONTENTS:

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY	1
INTRODUCTION	3
“A WARM BODY IS A WARM BODY”: TEACHER QUALITY POLICY UNTIL THE AUGHTS.	6
THE DATA EXPLOSION	12
ADDING VALUE TO THE DATA	14
THE TALENT MIND-SET: 2004 TO PRESENT	18
TNTP QUANTIFIES THE PROBLEM	18
RESEARCH AND POLICY COLLIDE—TEACHER EVALUATION TAKES CENTER STAGE ...	21
PREPARING AND COMPENSATING TEACHERS: WHERE WE ARE NOW	27
COMPENSATING TEACHERS BASED ON EFFECTIVENESS	29
PENSIONS	31
WHAT’S NEXT? RECOMMENDATIONS.....	32

ACKNOWLEDGEMENTS

The Joyce Foundation provided funding for this paper. The findings and conclusions are those of the authors alone and do not necessarily represent the opinions of the foundation. Likewise, the authors gratefully acknowledge the expertise and thoughtfulness of everyone who took time to talk with us about these issues, particularly Norman Atkins, Ann Best, Jim Blew, Celine Coggins, Tim Daly, Arne Duncan, Eric Hanushek, Kati Haycock, Jack Jennings, Brad Jupp, Jason Kamras, Joel Klein, Wendy Kopp, Jessica Levin, Gregory McGinity, Talia Milgrom-Elcott, Terry Moe, Linda Noonan, Van Schoales, Stu Silberman, and Kate Walsh.

DISCLOSURE

Numerous organizations and funders are mentioned in this paper. Some of these organizations have relationships with Bellwether Education or the authors. Bellwether counts the Bill & Melinda Gates Foundation, Carnegie Corporation of New York, and the Joyce Foundation among its funders. Bellwether works with the following organizations mentioned in the paper: Thomas B. Fordham Foundation, New Haven Public Schools, TNTP, Teach For America, Pittsburgh Public Schools, Teach Plus, Educators 4 Excellence, Hope Street Group, and America Achieves. In addition, Andrew J. Rotherham led the education project at the Progressive Policy Institute for eight years and was on the board of directors, and served as board chair, for the National Council on Teacher Quality. He is currently vice chair of the board of the Curry School of Education at the University of Virginia.

ABOUT THE AUTHORS

Andrew J. Rotherham is a Bellwether co-founder and partner and leads the Policy and Thought Leadership team at Bellwether Education Partners. He can be reached at andy@bellwethereducation.org.

Ashley LiBetti Mitchel is an analyst with the Policy and Thought Leadership team at Bellwether Education Partners. She can be reached at ashley.mitchel@bellwethereducation.org.

EXECUTIVE SUMMARY

Until recently, teacher quality was largely seen as a constant among education's sea of variables. Policy efforts to increase teacher quality emphasized the field as a whole instead of the individual: for instance, increased regulation, additional credentials, or a profession modeled after medicine and law. Even as research emerged showing how the quality of each classroom teacher was crucial to student achievement, much of the debate in American public education focused on everything except teacher quality. School systems treated one teacher much like any other, as long as they had the right credentials. Policy, too, treated teachers as if they were interchangeable parts, or "widgets."

The perception of teachers as widgets began to change in the late 1990s and early aughts as new organizations launched and policymakers and philanthropists began to concentrate on teacher effectiveness. Under the Obama administration, the pace of change quickened. Two ideas, bolstered by research, animated the policy community:

- 1) Teachers are the single most important in-school factor for student learning.
- 2) Traditional methods of measuring teacher quality have little to no bearing on actual student learning.

Using new data and research, school districts, states, and the federal government sought to change how teachers are trained, hired, staffed in schools, evaluated, and compensated. The result was an unprecedented amount of policy change that has, at once, driven noteworthy progress, revealed new problems to policymakers, and created problems of its own. Between 2009 and 2013, the number of states that require annual evaluations for all teachers increased from 15 to 28. The number of states that require teacher evaluations to include objective measures of student achievement nearly tripled, from 15 to 41; and the number of states that require student growth to be the preponderant criteria increased fivefold, from 4 to 20.

The high points are genuine breakthroughs. In Washington, D.C., a landmark new teacher evaluation system is improving the local teaching force. Elsewhere, in states across the country, new teacher evaluation systems are being implemented. Some of these new approaches are improving policy and practice; others are re-creating the shortcomings of earlier systems.

Still, much remains unchanged. Traditional teacher preparation programs still educate the majority of teacher candidates even as concern about the quality of preparation intensifies. Alternative route programs have sprouted and steadily grown in popularity, and data increasingly show that programs like Teach For America are reliable sources of teachers, but these programs are marginal relative to the overall number of teachers the country needs each year.

Teacher compensation based on effectiveness continues to capture the interest of policymakers, but the evidence about effectiveness and program design is mixed. Recent research on teacher pensions, a key part of how teachers are compensated, shows that states have been making it harder to qualify for a pension and decreasing benefits for new teachers in an attempt to address a \$390 billion pension funding shortfall. The result is a field that is less attractive to potential teachers.

To address these shortcomings and build on the momentum of the past 10 years, policymakers should consider five key issues:

- **You can't people-proof systems in education.** Current evaluation systems are a substantial improvement over previous policies. But are these the tools that will create a genuinely professional ethos for teachers? Evaluation systems should complement metric-driven systems with true managerial discretion. Districts should train and support managers and hold them accountable for their professional decisions.
- **Professionalize professional development.** The existing body of literature on professional development is extremely limited, but teachers must be supported in their work. Policymakers should identify and promote professional development that improves educator practice and student achievement. Evaluations should align with professional development for the purposes of growth and improvement, not just performance management.
- **Open and expand teacher preparation.** Teacher preparation is a difficult sector to reform, but doing so is key to improving teacher quality overall. Policymakers should increase rigor and quality in teacher preparation but also end protectionism of traditional preparation programs and open preparation to greater competition.
- **Address productivity.** Current education policy is often additive rather than productivity focused. Policymakers should find ways to promote productivity by better deploying the existing pool of teacher talent or improving how technology is used in schools and classrooms.
- **Address the politics.** Education is inherently political and the American debate about public education is special interest dominated. School improvement requires a robust political strategy to support its educational strategy.

INTRODUCTION

For years, the debate about American education was like a bad marriage. The arguments were about everything *but* the core issue—instructional quality. The other issues—education finance, school choice, standards—all matter, but are secondary to the importance of effective instruction. In the labor-intensive education field, effective instruction is nearly synonymous with teacher effectiveness. Trying to improve the quality of education in America without addressing teacher effectiveness is the same as trying to improve a baseball team without paying attention to hitting and fielding. Yet despite clear evidence about how much teachers matter, this is largely what American education tried for much of the 20th century.

That all changed quickly in the late 1990s and the aughts. Suddenly teacher quality emerged as a focus of national policymakers. New organizations launched and others made teacher quality a priority. The emphasis was so intense that it prompted a backlash, with some advocates decrying a “war on teachers.”¹

The pivot bemused researchers, who for decades had identified teacher effectiveness as the most important in-school factor affecting student achievement. And not surprisingly, as policymakers rushed to close the gap between research and practice, they made mistakes and overcorrections—the predictable and common problems of any significant public policy pivot. Except these policy changes affected America’s teachers. Teachers hold a conflicted place in American public life. They are at once individually beloved community figures tackling a difficult and important job and collectively among the most powerful interest groups in American politics. When policymakers began taking a serious look at teacher quality, the stage was set for a political battle that continues now at the national, state, and local level.

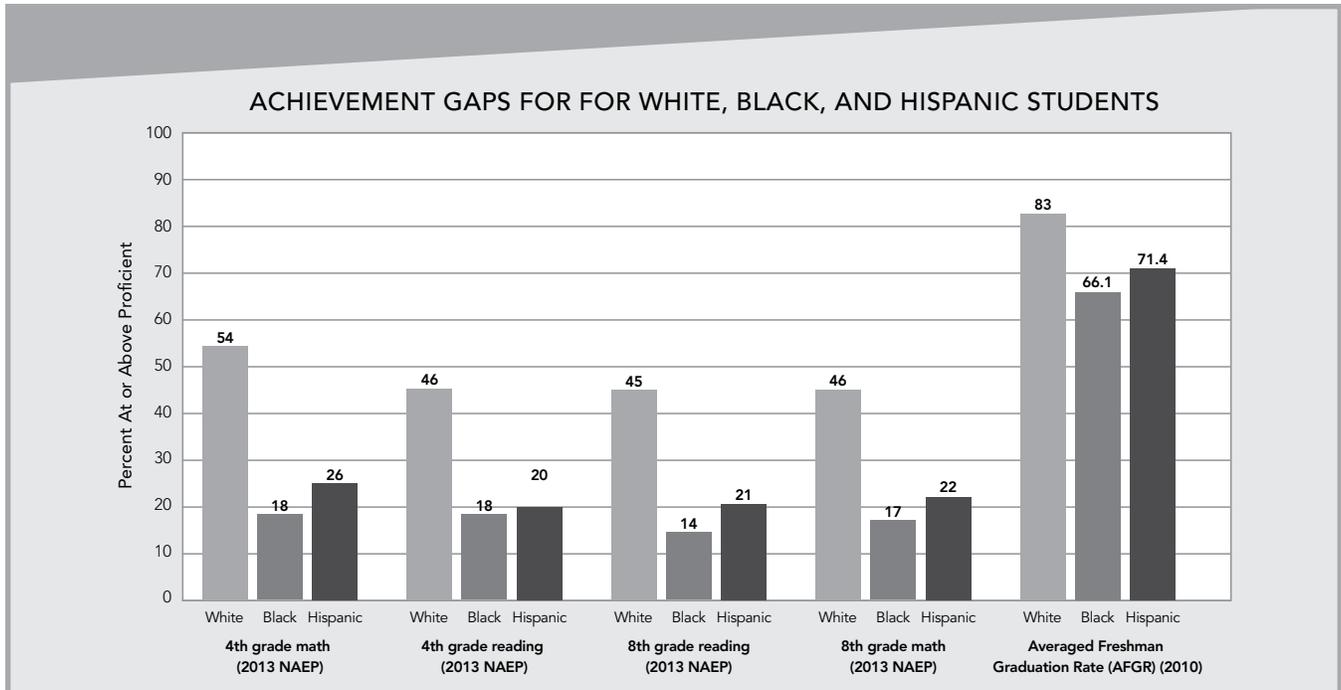
The story of this change is incomplete. It’s playing out in schools and statehouses around the country. It’s also full of puzzles and questions, some of which are beyond the scope of this paper. Because teacher effectiveness is so important, why did policymakers wait so long to take it on? If the answer is because teachers’ unions are so powerful, then why did change happen when it did—and under Democratic as well as Republican presidents and governors? What role did philanthropy play in driving these changes? Substantively, how much change has actually happened, or are we seeing old policy wine in shiny new bottles? Are the changes championed during the past few years likely to improve student learning? Are they even the optimal approaches for a field colliding with technology, evolving parental preferences, and a changing society?

This paper seeks to take an early look at some of those questions. It traces the changes since the late 1990s and attempts to capture a rapidly evolving status quo and make recommendations for next steps. It's based on the authors' research and analysis of existing literature, experience working directly on these issues in the public and nonprofit sectors, and interviews with experts, policymakers, researchers, philanthropists, and practitioners who played key roles leading teacher quality to where it is today.

The paper starts from several premises that underlie its narrative. First, with more than 3 million classroom teachers, effectiveness will vary—half will even be below average at any given point! Research shows clearly that effectiveness varies, is not highly correlated with common measures such as credentials or experience, and varies greatly within different schools. Second, as the legendary American Federation of Teachers President Albert Shanker rightly pointed out, with the number of teachers the country needs each year, some will be dreadful, just as there are dreadful doctors, lawyers, journalists, business leaders, managers, nurses, and practitioners of any field. The outliers shouldn't obscure all the teachers who work hard, care deeply, and change students' lives for the better.

But neither should our affection for teachers blind us to the challenges that do exist or to aggregate data that demand attention. Teachers themselves say some of their peers should go. In a 2008 national survey, 46 percent of teachers said they personally knew of one who is clearly ineffective and shouldn't be in the classroom. Sixty-three percent of teachers said they would strongly or somewhat support efforts to simplify the process for removing those ineffective teachers.² And principals agree. A 2006 survey showed that 72 percent of principals said that making it easier to fire ineffective teachers, including tenured teachers, would be a very effective way to improve overall educational leadership.³

Neither should our affection be used as a strategy or trump card to prevent policymakers from addressing these issues. In America today only 8 percent of low-income students receive a bachelor's degree by the time they are 24, compared with 82 percent of more affluent students.⁴ National Assessment of Educational Progress (NAEP) scores from 2013 show that white students are more than twice as likely as African American student to score proficient or above on reading and nearly three times as likely on math. The gaps in proficiency between white and Hispanic students are similarly alarming.⁵ While graduation rates are improving, schools still fail to prepare too many students for American life, and troubling outcome gaps persist: white students graduate high school at higher rates than their Hispanic and African American peers—83 percent, compared with 71.4 percent and 66.1 percent.⁶ These problems, and many others, cannot fairly be laid solely at the feet of teachers (or their unions). They are caused by many factors, but instructional effectiveness plays a role. Put plainly: Teachers matter and, as a result, policies about teacher quality matter a great deal. It's folly to pretend otherwise.



Source: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), *2013 Mathematics and Reading Assessments*, http://nationsreportcard.gov/reading_math_2013/#/student-groups; Robert Balfanz et al., “Building a Grad Nation”, February 2013, <http://www.americaspromise.org/sites/default/files/legacy/bodyfiles/BuildingAGradNation2013Full.pdf>.

Finally, this is a story about change at a contentious time in American education and American politics more generally. It’s a story about complicated and large-scale policy and practice changes. In just a few years, public policy shifted from not evaluating almost anyone in a meaningful manner to trying to evaluate millions of schoolteachers in ways that can differentiate teacher performance. It’s a tall order.

The resulting problems, debate, and false starts shouldn’t surprise anyone. The story can be read as a narrative of dysfunction, or it can be read as a narrative about needed but complicated change in a highly decentralized system. Fundamentally, despite the problems, it’s a story of progress. The last few years have produced real progress on teacher effectiveness and more generally on American schools, which, despite all the handwringing from the political left and right, have slowly been improving for several decades. And the focus on teacher effectiveness, while not always easy or comfortable for education leaders, has laid the foundation for a more genuine profession for teachers, which one can glimpse among all the activity and change.

“A WARM BODY IS A WARM BODY”: TEACHER QUALITY POLICY UNTIL THE AUGHTS

Teacher quality is hardly a new issue. How to train and credential teachers was a policy issue in the 19th century during the early days of public schooling. Individual school boards and locally elected citizens managed teacher training and licensing in an effort to control the quality of teachers in their communities. Gradually, training became more centralized as state officials and professional educators took over teacher licensure and certification.

By the middle of the 20th century, an agenda of professional standards sought to create a teaching “profession,” modeled after law and medicine. Proponents of professionalism pushed for reforms that increased the prestige of teaching. Among the proposals for reform were controlled entry to the field, increased state regulation, a national accrediting body, increased formal education requirements for certification, and higher salaries.⁷

Parts of this agenda influence education today. Yet opposition from within and outside the teaching field initially thwarted many of the goals of the professional agenda. That changed in the 1960s and 1970s as teachers’ unions, modeled after trade unions in their methods and ethos, gained traction by securing and using collective bargaining and other organizing strategies. Some of the first changes the unions went on strike for—better pay, working conditions, and respect for teachers—reflected the reforms proposed in the earlier professionalism agenda.

However, these reform efforts failed in their hope of building a genuine profession for teachers and in many ways cut against it. Commonality took precedence over professionalism. Teachers were trained and treated as interchangeable parts in the education system, more clerks than genuine professionals. This uniformity created political strength, power that improved schooling in some important ways. It also came with a cost for the field: rather than a professional ethos, it created a grievance culture. As one New York City teacher and teachers’ union district representative noted recently, “In too many schools teachers feel like well-paid assembly line workers.”⁸

At the same time a separate conversation was beginning among researchers—one of struggling to understand not what made teachers similar but what made them different. In 1966, *The Coleman Report* began to empirically reveal the importance of teacher quality for student achievement. The report found that socioeconomic factors affected student achievement most, but that teacher quality mattered more than all other in-school factors combined.⁹ The report set in motion efforts that continue to this day to discover which inputs matter to student outcomes. It has cast a shadow over American education for almost a half century. Some analysts and advocates read it as an

indictment of most reform efforts to hold schools accountable because of the effect of out-of-school factors. Others see it as a call to action; while addressing out-of-school issues involves a vexing mix of policy and political challenges, policymakers can take steps to address the within-school factors—especially teacher quality—now.

The 1983 Reagan administration report *A Nation at Risk* brought unprecedented attention to the problems in education and stoked the national debate about schools. Among other issues, *Nation* argued that the quality of the current teaching force was inadequate to the country's educational needs. The report cast a harsh light on the low caliber of teachers entering the workforce, the shortage of STEM educators, and the questionable quality of educator-preparation programs.¹⁰

A Nation at Risk had little immediate effect but sparked a conversation about improving schools that ultimately involved elected officials as well as business and education leaders. As Arkansas governor during the 1980s, before his presidency, Bill Clinton, for instance, proposed more stringent testing for teachers. In 1989, then-President George H.W. Bush and Clinton convened the nation's governors in Charlottesville, Va., to develop an agenda for raising standards and increasing accountability in education. States moved on their own, too. In 1993, the Massachusetts Education Reform Act (MERA) raised teacher certification standards, requiring new and veteran teachers to pass a subject-matter test for certification.¹¹

These efforts, however, were the exceptions. In the years leading up to No Child Left Behind and immediately following it, the teacher quality conversation was generally part of a larger discussion of broader school accountability reforms. To the extent that policy and philanthropy engaged with teachers specifically, it was to address quantity rather than quality. Many experts expected that increased retirement and student enrollment, fewer people entering education schools, and a policy focus on reduced class size would result in a massive teacher shortage. Behind the scenes the old adage that “there is never a teacher shortage on the first day of school” was more of a priority for policymakers than systemic improvement.

IS THERE A TEACHER SHORTAGE?

For the better part of the last three decades, analysts have speculated that a teacher shortage is looming. The *Chicago Tribune* warned of a shortage based on alarming statistics and predictions—in 1985,¹² in 2000,¹³ and again in 2002.¹⁴ Newspapers across the country, fueled by anecdotes from school districts and expert opinions, ran similar stories.¹⁵ Analysts projected that math and science courses, in particular, would be difficult to staff. Most stories cited baby boomer retirement, increasing student enrollment, and fewer candidates entering preparation programs as the causes. More recently, advocates argued that reform is driving teachers from the profession and creating shortages.

Yet the teacher shortage never quite happened, at least not as predicted. Public school enrollment increased annually through 2006, but so did the number of teachers. Between 1988 and 2001, the increase in teachers outpaced the increase in students, 29 percent to 19 percent.¹⁶ And college students continued to enter the teaching field. Between 2001 and 2005, the number of teachers prepared through traditional pathways increased 21 percent and the number of teachers prepared through alternative routes increased 104 percent.¹⁷

While shortages in math and science exist, the predictions were overstated. Analysts assumed that, in addition to increased student enrollment and teacher retirement, increases in course requirements would create a shortage of math and science teachers. Between 1987 and 2007, high school graduation requirements for math and science courses increased at a faster rate than any other core academic subject.¹⁸ But growth in the number of math and science teachers outpaced the growth in students. Enrollment in math courses increased by 69 percent, but the number of math teachers increased by 74 percent. Similarly, enrollment in science courses increased by 60 percent, but the number of science teachers increased by 86 percent.¹⁹ As of 2004, 80 percent of high school teachers were certified in their main assignment area.²⁰

And in some states, a surplus, rather than a shortage, has been the larger concern in recent years. Data show that nearly a dozen states overproduced elementary school teachers in 2010. Delaware, Michigan, New York, and Pennsylvania all produced at least 200 percent of the necessary number of elementary teachers. Illinois is one of the worst offenders: the state produced 9,982 elementary teachers for 1,073 positions that year.²¹ Illinois teachers also had difficulty finding jobs in traditionally hard-to-staff subjects like math, science, and special education.²² Glenview School District 34, for example, received 4,300 applications for 74 positions in 2009, and the number of applications to Chicago Public Schools doubled between 2003 and 2008.²³

A closer look at retirement data indicates that there, too, the rhetoric is overblown. Between 1988 and 2008, retirements increased from 35,000 to 85,000, and the number of teachers over the age of 50 doubled, from 530,000 to 1.3 million. Yet currently, the modal age of teachers is 59, suggesting teacher retirements should be at an all-time high, but there were 2,000 fewer retirees between 2004 and 2008, and between 2008 and 2012 there were 250,000 fewer teachers over the age of 50. Data also show that teacher retirements have consistently accounted for about a third of teachers leaving teaching, and only 14 percent if transfers are included.²⁴ Overall, the number of continuing teachers hovered between 87 percent and 86 percent between 1987 and 2008.²⁵

Although predictions of an overall shortage were off base, states, districts, and schools consistently struggle with shortages in particular disciplines and communities. High-poverty schools are among the hardest hit. In every subject area, schools with a higher concentration of students who receive free or reduced-price lunches have fewer teachers certified in their assignment area. Only 54.5 percent of math teachers in high-poverty schools are certified in math, compared with 80.3 percent in low-poverty schools.²⁶ Certain states also have difficulty finding teachers with the right certifications. In Louisiana, Delaware, Washington, and Nevada, over a quarter of core academic secondary classes in 2007 were taught by teachers without the appropriate content-area major or certification.²⁷

There's also a shortage of bilingual and special education teachers. In 2007, some 28 percent of bilingual education teachers in New York were not certified, the majority of whom taught in New York City.²⁸ In Florida, 30.5 percent of bilingual education teachers hired in 2010 did not have the appropriate certification.²⁹ Some Texas districts host recruitment fairs across the border in Mexico to fill their bilingual education gaps.³⁰

In the early part of the 2000s, the shortage of special education teachers was at its height. The number of uncertified special education teachers increased 23 percent from 1999 to 2001.³¹ Since then the shortage has declined, but the problem persists. In 2010, every state except Oklahoma reported a shortage in special education teachers.³² Between 2006 and 2011, the number of uncertified special education teachers decreased 61 percent, from 49,058 to 19,242.³³ While this is a positive trend, many students with disabilities are at a disadvantage. Class size limits vary depending on the district and type of disability, but many districts, such as Chicago Public Schools, allow between 15 and 17 students with disabilities per class period.³⁴ That means that in 2011, we can estimate that between 288,630 and 327,114 students with disabilities were taught by an uncertified teacher. Not a national teacher shortage, but an acute problem nonetheless.

In 1987, the National Board for Professional Teaching Standards was established to provide a credential to advanced or “master” teachers, but generally there was little effort to reform the routes into the profession used by most teachers. Later, the American Board for Certification of Teacher Excellence (ABCTE) launched a virtual accrediting program to give teachers another route to a portable teacher credential. Both of these initiatives remained marginal in different ways. NBPTS certified thousands of teachers, but studies showed that teachers with the credential were only slightly more effective, at best, than others.³⁵ ABCTE, meanwhile, was adopted in few states and failed to create the widely used portable national credential its founders imagined.

After a great deal of debate and a change in partisan control of the White House, the ideas Bush and Clinton introduced in Charlottesville came to fruition with the Improving America’s Schools Act of 1994 (IASA)—a revision of the Elementary and Secondary Education Act (ESEA). IASA required states to develop content-area standards and corresponding assessments and describe “adequate yearly progress,” based on performance on those assessments, as a condition of receiving federal Title I funding.³⁶

During the 1990s, however, analysts and civil rights leaders became increasingly concerned that states were evading and muting the intended effects of these reforms. They worried that the problem of inequitable access to highly effective teachers was persisting, if not growing. Equity-oriented education reformers struggled to build alliances across ideological lines. Their positions did not align with traditional Republican or Democratic ideas on education. As the decade wore on an informal coalition began to emerge and liberal Capitol Hill Democrats such as Dale Kildee (Mich.) and George Miller (Calif.) put forward ideas to increase accountability and require states to increase equity in the distribution of teachers.

“Ten years ago, hiring teachers was literally viewed as an operational matter. It was the equivalent of just filling a vacancy.”

– Wendy Kopp, founder of Teach for America

These ideas ultimately found their way into the No Child Left Behind Act of 2001 (NCLB). In particular, NCLB added an accountability measure for teachers, requiring that all teachers in core content areas be “highly qualified” by the 2005 school year based on state-developed definitions of quality.³⁷ Substantively, the provisions became a paper chase as states, under intense pressure from the teachers’ unions, sought ways to minimize the impact of the rules on veteran teachers. Politically, however, they stimulated increased attention to teacher quality and set the stage for more ambitious federal policy efforts to improve teacher effectiveness.

In the early part of the 2000s, national organizations began to examine the human capital problem in education with more intensity. The Thomas B. Fordham Foundation, which had long championed alternative certification and merit-based pay schemes, launched the National Council on Teacher Quality, an independent organization intended to promote a range of human capital reforms. Since then, NCTQ has grown into a respected national advocacy organization. At the same time the Progressive Policy Institute (PPI) released several controversial papers, by analysts such as Bryan Hassel and Rick Hess, challenging policymakers to overhaul how teachers were credentialed and paid. These ideas would later become more accepted but at the time were highly contentious even within the policy community—especially coming from the left-leaning PPI.

“When I first started as Chancellor, people were expendable. A warm body was a warm body, and no one was worried about having the right people in the right places.”

– Joel Klein, former Chancellor of New York City Department of Education

A shift was on and although it had not permeated the states, at the elite level of national policy a consensus was emerging: Neither prospective nor veteran teachers were interchangeable as long as their credentials were in order. “Ten years ago, hiring teachers was literally viewed as an operational matter,” Wendy Kopp, founder of Teach For America, says. “It was the equivalent of just filling a vacancy.”³⁸ Joel Klein, who served as chancellor of New York City’s public schools for eight years, observed the shift firsthand. “When I first started as chancellor,” Klein said, “people were expendable. A warm body was a warm body, and no one was worried about having the right people in the right places.”³⁹

THE DATA EXPLOSION

In the decades following *The Coleman Report*, higher quality administrative data allowed for quantitative research on teacher effectiveness with more specificity and rigor. Two key studies—one conducted by Eric Hanushek, John Kain, and Steven Rivkin (1998); and another done by Dan Goldhaber and Dominic Brewer (1999)—found that the role of the teacher has greater predictive power than any other in-school factor, including teacher training and certification. Hanushek, Kain, and Rivkin used school- and student-level administrative data from the Texas Education Agency to examine the effects of different variables on student achievement. They found that variation in teacher quality explains at least 7.5 percent of the variation in student achievement, with “reasons to believe the true percentage is much larger.”⁴⁰ Goldhaber and Brewer used *National Education Longitudinal Study* survey data from 1988 to conduct a similar analysis, ultimately estimating that teacher effectiveness explained 8.5 percent of student achievement.⁴¹

These analyses did not stay isolated in the academic community but were blasted into the policy and political communities. The early results were awkward political compromises. The No Child Left Behind Act, for instance, required that all teachers be “highly qualified” rather than emergency credentialed. The law largely left it to states to define what “highly qualified” or HQT requirements looked like in practice.⁴²

“It took districts three or four years to prove that HQT was a minimum standard, but most people realized that minimum was a bad standard.”

– Brad Jupp, advisor on teacher quality,
U.S. Department of Education

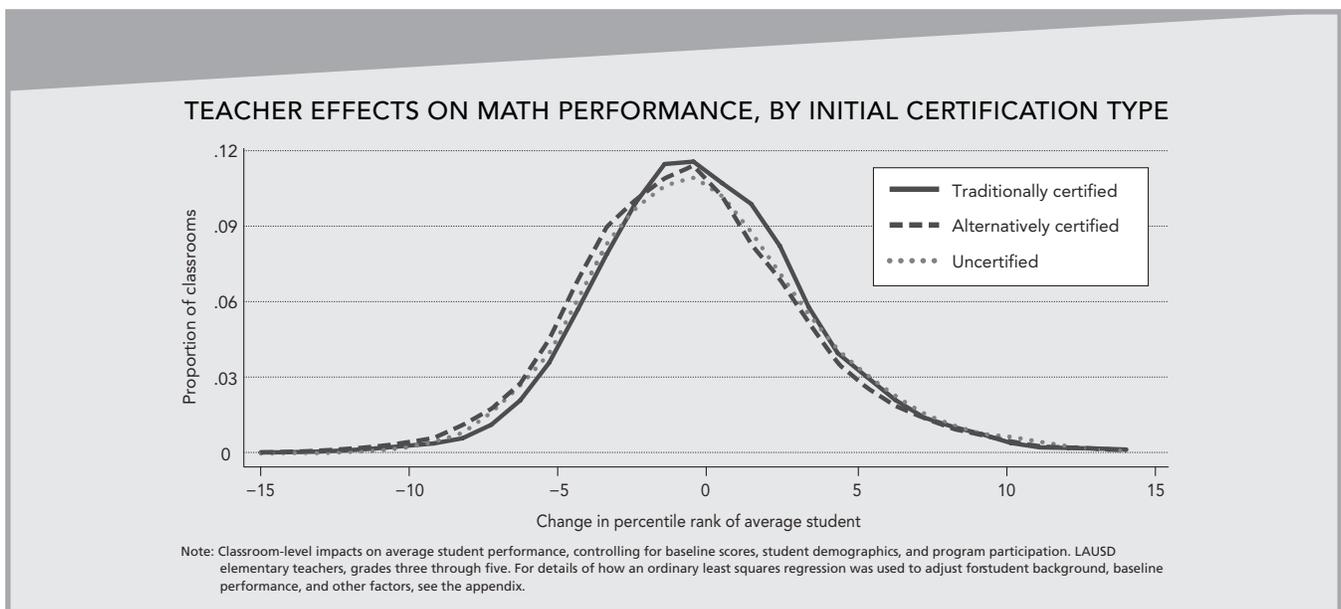
Initially, the HQT provisions were not intended to improve overall teacher effectiveness. They were explicitly billed as a way to ensure equitable distribution of effective teachers, particularly in low-income and minority communities and schools. The Education Trust advocated for HQT as an equity measure. Representative George Miller, one of the four architects of NCLB, was also a proponent of the equity potential NCLB offered. In practice, because of the way states chose to implement the law, the HQT provisions ensured neither equitable distribution nor teacher effectiveness. Because of intense pressure from groups representing teachers, states implemented the law in a way

that ensured that veteran teachers would not lose their jobs because of the new requirements. The resulting approaches focused on compliance and paper chasing rather than on actual measures of what teachers know and can do. One could achieve HQT status by attending conferences or other weak measures instead of demonstrating subject matter mastery. Within a few years almost everyone was “highly qualified” on paper. As Brad Jupp, one of Secretary of Education Arne Duncan’s key policy advisors on teacher quality, says, “It took districts three or four years to prove that HQT was a minimum standard, but by then most people realized that minimum was a bad standard.”⁴³

The HQT provisions highlighted the schism that underlies most policy and political debates about teacher quality. The teachers’ unions and their constellation of interest groups can hardly resist the evidence about the importance of teachers. After all, the importance of teachers is their bread-

and-butter advocacy point. But when that importance translates into accountability systems or other consequential measures that can carry adverse consequences for low-performing teachers, these same advocates find themselves in an awkward pincer. They're left arguing that teachers are important, even heroic, in their work but not so much so that they should be held accountable like other professionals. In other words teachers are important—except when they're not.

Those hoping that discrediting HQT would bring back the previous status quo of inattention to teacher quality, or otherwise put the human capital genie back in its bottle, were disappointed. Instead, policymakers began searching for new ways to evaluate teacher effectiveness. In 2006, for example, Robert Gordon, Thomas Kane, and Douglas Staiger published *Identifying Effective Teachers Using Performance on the Job* as part of the Hamilton Project at the Brookings Institution, a highly regarded Washington think tank. In the paper, Gordon, Kane, and Staiger analyzed the performance of approximately 150,000 students in 9,400 Los Angeles Unified School District classrooms for each year between 2000 and 2003. They found that there was no statistically significant difference in the effect on student achievement between certified, alternative route, and non-certified teachers.⁴⁴ What the paper said, and where it was published—Brookings, a middle-of-the-road organization—called into question the credential-based approach to teacher quality, raised serious questions about the teacher tenure process, and further thrust value-added analysis into the policy mainstream.



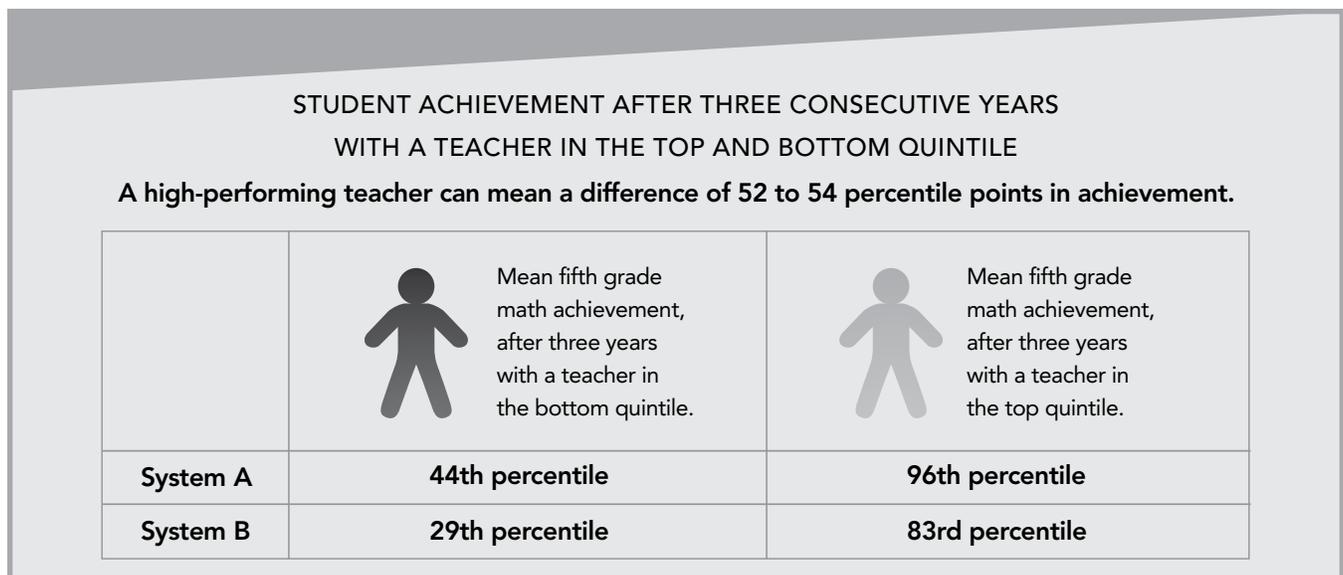
Source: Robert Gordon, Thomas Kane, and Douglas Staiger, *Identifying Effective Teachers Using Performance on the Job*, Brookings Institution, 2006, http://www.brookings.edu/views/papers/200604hamilton_1.pdf.

ADDING VALUE TO THE DATA

Actual data, especially improved administrative data, on student outcomes presented policymakers and advocates with a compelling alternative to using training, experience, and other credentials to measure teacher effectiveness. The idea was not a new one.

As early as 1971, Eric Hanushek released research on using student outcomes to evaluate teacher effectiveness.⁴⁵ Later, William Sanders and Robert McLean, at the University of Tennessee, became known for their work developing “value-added” models. Sanders and McLean experimented with statistical methods using administrative student data starting in the early 1980s; their work developed to the point that, in 1991, Tennessee passed the Education Improvement Act, which relied heavily on their research. Their value-added methods or models use multiple years of data to evaluate teacher effectiveness. Sanders and McLean’s research showed that the variation in quality between teachers is significant, and the effects on student achievement are lasting and cumulative: a student taught by teachers in the top quintile for three consecutive years scored between 52 and 54 percentile points higher than a student who started at the same proficiency level but was taught by teachers in the bottom quintile for the same period.⁴⁶

Kevin Carey was one of the first analysts to suggest using value-added data aggressively as a policy strategy. In 2004, he released a paper with the Education Trust recommending that policymakers consider using value-added methods to evaluate teacher quality. Gordon, Kane, and Staiger then published their provocative Brookings paper that, in addition to revealing issues with traditional measures of teacher quality, recommended policy and practice changes that would: reduce the



Source: William L. Sanders and June C. Rivers, “Research Progress Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement,” November 1996, http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf.

barriers to entry to teaching, reward teachers who improve student achievement, and dismiss the bottom quartile of teachers after two years of teaching. Collectively, the attention from the political left to these issues changed the debate. Van Schoales, a leader on efforts to improve teacher quality in Denver, says these analyses, “provided a place for policy folks to have a conversation that was critical. We wouldn’t be where we are now without it.”⁴⁷

No Child Left Behind’s impact on teacher quality largely came from issues unrelated to the HQT rules. Instead, the law required states to assess students annually in grades 3-8, producing more consistent student achievement data, first reported by states at the end of the 2005-2006 school year. The result was a boon for policymakers and researchers seeking to use value-added methods.⁴⁸

Value-added approaches became highly controversial as they were employed as a strategy to inform evaluations or award tenure. Unions, never warm to the value-added idea, are now staunchly opposed to including value-added data in evaluation systems. In late 2013, Randi Weingarten, president of the American Federation of Teachers, reversed her previous support for some use of value-added and attacked evaluation systems by referring to value-added methods as “black-box algorithms” that are “incomprehensible, at least to those who don’t have a Ph.D. in advanced statistics.”⁴⁹

In practice, value-added data comprise just one component of any evaluation system, and cover only about a third of teachers because most teach in subjects and grades that are not assessed. States require observations of teacher practice more often than other measures in their teacher evaluations. In 2013, 45 states and D.C. required observations, compared with 17 states that require parent or student surveys and 20 states that require student growth as the preponderant measure.⁵⁰ Evaluations by RAND,⁵¹ The Gates Foundation’s MET project,⁵² and other researchers show that while value-added models should be used with caution, they can help responsibly inform some personnel decisions and are not the lottery their critics make them out to be. Yet critics have seized on technical elements of value-added data as a way to undercut teacher evaluations generally. Rushed implementation, poor management capacity in school districts, and high-profile mistakes gave critics plenty of talking points and fuel. When the *Los Angeles Times*, and subsequently other newspapers, published value-added data linked to individual teachers, that threw fuel on the fire and furthered a perception that value-added was all about shaming teachers.

All of the critiques and cautions are not without merit. At its core, however, much of the opposition to evaluation systems is actually more fundamental opposition to linking evaluations to consequential personnel decisions that can result in veteran teachers losing their jobs. With a few exceptions, measures with that effect have proven a bridge too far for the teachers’ unions.

NEW HAVEN: UNION COLLABORATION

In 2009, New Haven Public Schools signed a contract that school reformers and teachers' union leaders lauded as a national model. The contract granted teachers annual raises and explicitly allowed districts to tie student performance to teacher evaluations and bonuses, laying the groundwork for the dismissal of low-performing teachers in New Haven. Yet New Haven's union support and ethos of collaboration appeared in stark contrast to similar reforms in Washington, D.C.

Some of this was political—stung by bad publicity in D.C., the American Federation of Teachers needed to show it could agree to something of consequence. Some was contextual—New Haven officials were prepared to take steps to reform evaluation with or without the union's involvement. Still, the collaborative process produced a positive outcome, and New Haven teachers ratified the contract with a vote of 842-39 only a few weeks after teachers and students protested the dismissal of 229 DCPS teachers.

The New Haven evaluation model rates teacher performance based on student learning outcomes, instructional practice, and professional values. Each teacher has a mid-year and end-of-year conference. Teachers who are on track to be rated Exemplary or Needs Improvement, the ends of the spectrum, have their ratings confirmed by a union-approved "validator." Teachers rated Needs Improvement immediately receive "immediate and intense" professional development. If tenured teachers receive Developing ratings two years in a row, they are treated as if they received a Needs Improvement rating.⁵³

Since the 2009 contract, New Haven teachers have received three years of evaluation ratings. In those three years, nearly 5 percent of teachers have been pushed out of their positions. Yet not every teacher who received a Needs Improvement rating left the district. In the first year, 75 teachers were identified as Needs Improvement; 29 of those teachers improved their performance enough to keep their jobs and 8 teachers left the district before receiving their summative rating. Thirty-four of the remaining low-performing teachers—16 tenured and 18 non-tenured—resigned or retired.⁵⁴ The rest were permitted to keep their positions. In the second year, 58 teachers were identified as low-performing. Only 28 were pushed out; 20 increased their rating during the year and 10 were granted another year to improve.⁵⁵ Last year, 36 teachers were at risk of dismissal; 16 improved their rating during the school year and 20 were pushed out.⁵⁶

Unlike in other districts, the union did not dispute any of the evaluation ratings. To date, all teachers who have been pushed out have been persuaded to resign by the district rather than dismissed. It's unclear if this practice is sustainable, but it seems possible. The union role in selecting a validator and the opportunity to improve during the school year are substantial motivators and reduce the surprise that sometimes goes with poor evaluation ratings.

The 2014 contract—ratified 775-79 late last November—builds on the union-district collaboration in New Haven. It also builds on the 2009 reform by adding a monetary component. Under the new contract, teachers who work in hard-to-serve schools can receive extra pay, different work rules, and additional training; teachers rated as Needs Improvement or Developing cannot receive automatic raises unless they attend extra training sessions; and teachers rated Effective or above are eligible for leadership roles as teacher "facilitators."

There is also a Lucy-and-the-football quality to the evaluation debate, with reformers playing the part of Charlie Brown. Teachers' union leaders repeatedly promised that with better standards and evaluation systems they would be open to consequential evaluations to address low-quality instruction. With those elements on the horizon, suddenly a new set of issues emerged as the key ones that must be addressed before they can take this step. Currently implementation of the new Common Core State Standards is the obstacle that must be addressed before evaluations can proceed.

The unions are being squeezed between two different concerns. They are uncomfortable with evaluations based on managerial discretion—and not without some reason, given the state of education management. And they are also uncomfortable with evaluations based largely on outputs in the classroom—again a position that is not entirely without merit. They propose half-measures such as “peer review” plans for new teachers, but while these steps help at the margins they fail to substantially address the problem. As one longtime teachers' union leader confided, peer review can and does address observably bad teaching—those teachers who cannot manage a classroom, organize and sequence teaching and learning, or arrive at work prepared and in a condition to work. It is not an effective strategy to address the more widespread problem of what might be called palliative teaching—classrooms that are seemingly well run, where students are not at immediate risk, but where they are learning little.⁵⁷ Absent a compelling alternative, the unions' position is fundamentally defensive and, over time, untenable despite their current political prowess.

THE TALENT MIND-SET: 2004 TO PRESENT

Although the research pointing to teachers as a key leverage point in education wasn't new, it wasn't until the early 2000s that it took hold with policy leaders. "Around [2004] was really when the existing evidence became accepted," says Linda Noonan, of the Massachusetts Business Alliance for Education. "The teacher is the greatest factor in the classroom, and an effective teacher could move a student farther than an ineffective teacher."⁵⁸ Policy, and politics, began to catch up to the research.

TNTP QUANTIFIES THE PROBLEM

Across the education field, there was a general acknowledgement of the problems many existing policies caused and how they hampered school effectiveness. Politics, internal teachers' union disagreements, a lack of clear alternatives in some cases, and, perhaps most importantly, a lack of quantifiable analyses of these issues kept the conversation in the background. The New Teacher Project (subsequently rebranded as just "TNTP") set itself to the work of documenting the actual impact of these issues. It is hard to overstate the impact of its 17 years of work.

TNTP, a nonprofit that spun out of Teach For America in 1997, used existing data to quantify the effects of various policies and practices. The organization published research using actual data to show just how little school districts took effectiveness into account in various decisions about teachers. TNTP's research on district hiring and staffing practices—particularly its seminal report *The Widget Effect*—was pivotal to advancing the teacher effectiveness agenda.

"Mutual consent [between schools and teachers] has to be put in place. Without it, low-income schools become the dumping ground for poor performers."

– Jim Blew, program officer,
The Walton Family Foundation

TNTP's first reports, *Missed Opportunities* and *Unintended Consequences*, brought much-needed attention to dysfunctional district hiring and staffing practices and how they hurt students in high-need schools.⁵⁹ *Missed Opportunities* showed that urban districts' hiring practices—in part attributable to contract requirements—resulted in less effective candidates.⁶⁰ *Unintended Consequences* revealed related problems. Common provisions in teachers' contracts require school leaders to prioritize voluntary transfers (veteran teachers who want to move between schools in a district) and excessed teachers (teachers who were cut from a school in response to declines in budget) over all other applicants in staffing decisions, regardless

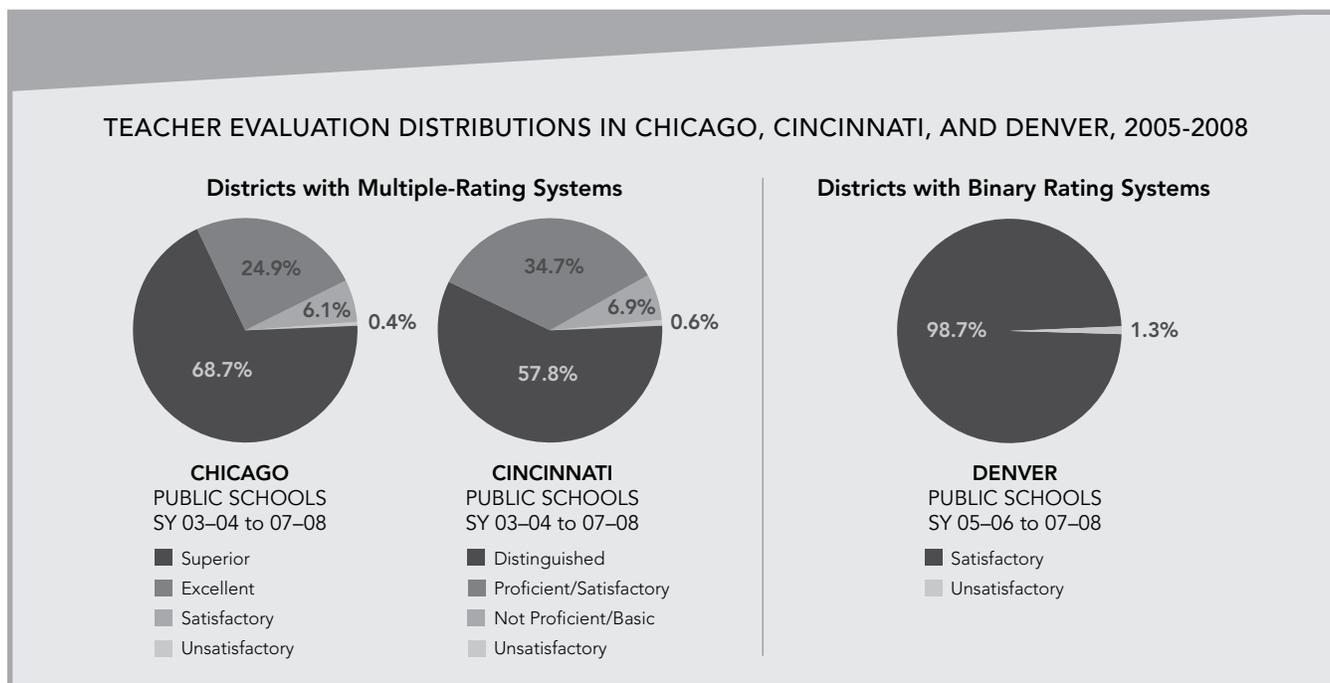
of fit, qualifications, or leadership preferences. Consequently, many schools have little or no choice in the teachers they receive. In practice the voluntary transfer and excessed teacher processes often serve as an easier alternative to termination, so ineffective teachers are passed around as forced placements. As a result, ineffective teachers are often pushed to high-need schools, undermining efforts to improve the quality of teachers for high-need students. “Mutual consent [between schools and teachers] has to be put in place,” says Jim Blew, of the Walton Family Foundation. “Without it, low-income schools become the dumping ground for poor performers.”⁶¹

The evidence from these reports led districts such as New York City and the District of Columbia to pursue reforms to the hiring and staffing requirements in their contracts. In 2005, then-Chancellor Joel Klein was battling with the United Federation of Teachers (UFT) over hiring provisions in the union contract. The work rules he fought for were, among others, to end forced placements of teachers against the will of principals and instead ensure mutual consent for teachers and principals. TNTP’s *Missed Opportunities* and *Unintended Consequences*—reports that empirically revealed how district staffing practices, as stipulated in union contracts, undermine the quality of urban schools—drove much of Klein’s focus. The two sides could not agree and the issue ended up at an arbitration hearing: the last chance to convince a panel of three arbiters of how the contract should read.

At that point, the arbiters most often sided with UFT in such hearings. “Generally, fact-finding got [the UFT] what they wanted,” Klein says.⁶² This time, Klein brought in then-President of TNTP Michelle Rhee to testify at the hearing, and she brought ample data to support her testimony. She produced data on the number of forced placements in the district, the school sites that received the most forced-placement teachers, and the effect of forced placements on students. Data like these had not previously been part of the process, and one observer in the room described the UFT’s representatives as “stunned.” The arbiters ruled primarily in Klein’s favor, ending forced placement and requiring mutual consent.

That 2005 arbitration hearing, little noticed outside of New York except by experts, had lasting significance for two reasons. It was one of the first times a district had used data to win an arbitration hearing and signaled a new era in which data, rather than assertions about various work rules, could carry more weight. Second, it was one of the first times that a district and union agreed to a contract with those staffing rules, which set the foundation for future contracts such as those in Washington, D.C., in 2010 and elsewhere.

TNTP's *The Widget Effect*, released in 2009, demonstrated the failure of existing teacher evaluation systems to differentiate teacher performance. The report revealed serious flaws in the way districts, ranging in size from 4,450 to 431,700 students, evaluate teachers. For example, between 94 and 99 percent of the 15,000 teachers in TNTP's sample were rated good or great in their most recent evaluation, despite their students continuously failing to meet basic academic standards. In Chicago, only 0.4 percent of teachers were rated Unsatisfactory between 2004 and 2008,⁶³ despite well-documented problems in that school system. In Denver, only about one in six schools not meeting performance goals in 2008 had a teacher rated Unsatisfactory. None of Cincinnati's schools identified as persistently low-performing had a single tenured teacher rated Unsatisfactory during the same time.⁶⁴ In the aggregate, public schools were systematically failing to manage the performance of the most important part of the educational chain.



Source: David Weisberg et al., The New Teacher Project, "The Widget Effect," <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.

Instead, teachers with consistently poor performance were not remediated or counseled out of their position. Teachers with consistently high performance were neither rewarded nor recognized. The evidence suggests this pattern is widespread and, by failing to reliably differentiate teacher performance, districts' evaluation systems were effectively treating teachers as uniform, interchangeable parts, or widgets.⁶⁵

The Widget Effect had national impact. It shifted the focus to teacher evaluation systems as the key lever to improve teacher quality. The authors of *The Widget Effect* offered recommendations for making teacher evaluation systems more rigorous and meaningful, the first of which was to evaluate and differentiate teachers based on their effect on student performance. The authors included a sidebar on value-added methods, suggesting their promise as an addition to and component of comprehensive teacher evaluation systems, not a stand-alone reform.⁶⁶ This position was wildly mischaracterized by critics of TNTP and value-added evaluation in general, who claimed teachers would be fired based on value-added data. In practice, no state or school district has used value-added data as the sole criterion in any evaluation system. It is indicative of the confusion—and the deliberate misinformation—that characterizes the debate about teacher evaluation.

RESEARCH AND POLICY COLLIDE—TEACHER EVALUATION TAKES CENTER STAGE

In 2009, research, policy, and practice, as well as the growing concern about teacher effectiveness, collided at the national level and in high-profile cities.

In Washington, D.C., controversial school chancellor Michelle Rhee introduced IMPACT. The IMPACT system evaluated teachers on their impact on student achievement, instructional expertise, collaboration, and professionalism.⁶⁷ IMPACT weighted teachers' effect on student achievement above other factors: initially, 50 percent of a teacher's rating was based on value-added data and 40 percent was based on observational ratings.⁶⁸ Since then, IMPACT has been revised; currently, 35 percent is tied to value-added methods and 15 percent tied to student learning goals developed by teachers.⁶⁹

WASHINGTON, D.C.: EVALUATION AND COMPENSATION

Washington, D.C.'s evaluation and compensation systems, IMPACT and *IMPACTplus*, are two of the most highly regarded teacher quality reform efforts in the country. Then-Chancellor Michelle Rhee introduced IMPACT in the fall of 2009. *IMPACTplus* was instituted as part of the 2010 teachers' contract.

IMPACT represented one of the most ambitious attempts to translate the research about teacher effectiveness and student learning into practice in a district. Jason Kamras, chief of Human Capital at DCPS and former National Teacher of the Year, worked alongside Rhee on IMPACT and *IMPACTplus*. "In doing the research leading up to IMPACT, what we generally found was that there was good information about the idea [of evaluation], but almost no information on how to do it," Kamras says. "There were value-added models out there, for example, but no one was figuring out how to take a value-added model and turn it into an evaluation system."⁷⁰

IMPACT is an evaluation system that rates teacher performance based on student achievement data, instructional expertise, and collaboration within the school community. Student achievement data are either value-added data from standardized tests or student performance data from teacher-delivered assessments.⁷¹

The early evidence on IMPACT is promising. Research from the University of Virginia found IMPACT may improve teacher performance and retention: In the first three years, mean teacher scores improved by 10 points; and teachers who received base-pay increases were more likely to continue teaching than those who did not. Washington, D.C.'s NAEP scores suggest IMPACT may be helping drive improvements in student achievement. Compared with 2009, when IMPACT was instituted, the 8th-grade math performance on the 2013 NAEP increased by 12 points, the fastest gains in the country. The city also has made notable gains in 8th-grade reading and 4th-grade math and reading since 2009.⁷² There are likely multiple reasons for those gains, including IMPACT, but at a minimum it's difficult to argue (although some do) that the new system is hurting student learning.

IMPACT may also reduce the number of low-performing teachers. In 2008, before IMPACT, zero percent of teachers were rated Ineffective. In 2009, 2 percent of teachers were rated Ineffective. Some 113 teachers—over 90 percent of all Ineffective teachers—were dismissed because of Ineffective ratings. Further, the threat of dismissal was often enough to push low-performing teachers out; teachers who were rated Minimally Effective closer to the Ineffective threshold were more likely to voluntarily exit than those rated Minimally Effective closer to the Effective threshold.⁷³

At the same time, IMPACT and *IMPACTplus* help recognize high-performing teachers. Teachers rated Highly Effective can earn a base salary increase of up to \$25,000.⁷⁴ And local philanthropic groups recognized Highly Effective teachers with accolades like the Standing Ovation for DC Teachers ceremony and the Excellence in Teaching awards. Effective teachers, not surprisingly, comprise a much larger group than teachers rated Ineffective, showing the extent to which the problems caused by low-performers obscure the great work of the many high performers.

IMPACT was notable as a national model for several reasons. First, it was an ambitious and comprehensive evaluation system for all teaching personnel in Washington, not just those in core subjects. In addition, it was the first large-scale effort to implement a teacher evaluation system tying personnel decisions to student achievement outcomes. It was also noteworthy because school officials sought to explicitly link the evaluation system to the contract so that teachers could be dismissed for poor performance. Teachers would be able to file a grievance if they didn't have a proper evaluation but would not be able to simply argue against an evaluation outcome they just didn't like or agree with. It was an unprecedented amount of discretion for a district to have on decisions affecting its teaching force.

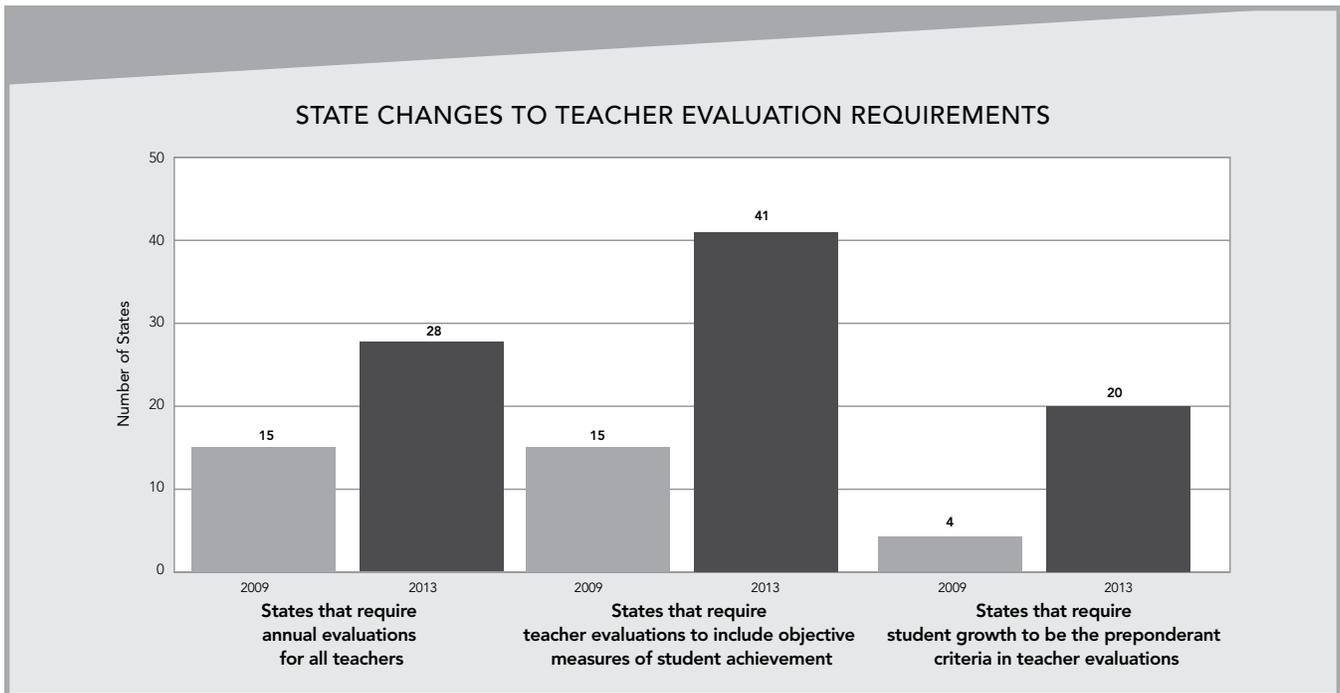
Under the new contract, in addition to requiring “mutual consent” of the teacher and the school for any teacher placement when layoffs occur, layoffs are based on performance in the classroom instead of simply when someone was hired. Washington Teachers' Union (WTU) and District of Columbia Public Schools (DCPS) also agreed to a performance-pay system, *IMPACTplus*, as part of the contract: teachers rated Highly Effective under IMPACT are eligible to enroll in a performance-pay system, under which they can receive significant pay raises or be more easily dismissed.⁷⁵

Not surprisingly, development and implementation of IMPACT created intense friction between WTU (and its parent organization, the American Federation of Teachers) and DCPS, particularly Chancellor Michelle Rhee. Both sides saw that IMPACT would be precedent setting, but only the schools saw that precedent as a worthy one. In 2010, after contentious debate, both parties agreed and WTU members ratified the contract with a vote of 1,412 to 425.⁷⁶ The outcome points up two issues. After nearly a year of intense and public negotiations between district and union leadership, actual teachers favored the contract 3 to 1—further demonstrating the disconnect between public narrative and reality. At the same time, approximately 4,000 teachers were eligible to vote in the contract election, but less than 50 percent of teachers did. The turnout rate speaks to the lack of engagement of many teachers in issues affecting their profession. As a rule, nonvoters far outpace voters in most teachers' union elections, even high-profile ones.

Against this backdrop, the U.S. Department of Education was introducing the first round of Race to the Top (RTT), a \$4.35 billion competitive grant program funded through the American Recovery and Reinvestment Act (ARRA) of 2009, commonly referred to as the Stimulus or Recovery Act. RTT had multiple goals and priorities, all aimed at improving student achievement. Yet teacher and principal quality was a core issue—the “Great Teachers and Leaders” section was weighted more heavily than any other component, comprising 28 percent of the total application points.⁷⁷ RTT has been attacked as the administration's effort to force its charter school policies on states. In fact, charter schools and all other innovative public schools were only worth 40 points, or 8 percent, of the competition's 500 winnable points. Teacher evaluation was the thrust, and states responded.

Between 2009 and 2013, the number of states that required annual evaluations for all teachers nearly doubled, from 15 to 28. The number of states that require teacher evaluations to include objective measures of student achievement increased from 15 to 41, and the number of states that require student growth to be the preponderant criteria increased from 4 to 20.⁷⁸ Winning states are struggling to implement many aspects of their RTT plans, including the evaluation component.⁷⁹ However, it seems the longer-term impact will be more than the specific actions of states that won the competition. RTT sparked a sea change in state policy.

RTT created a rare moment of bipartisan alignment. Governor Rick Scott, for example, pushed a Republican-led legislature to pass Florida’s SB 736 in March 2011, after former Governor Charlie Crist—at that point a Republican, he later became a Democrat—had vetoed a similar bill the previous year. Indiana’s Republican legislature and Governor Mitch Daniels passed an ambitious evaluation law in April 2011, as did Michigan in July 2011. In addition, Republican governors led successful evaluation legislation efforts in Idaho, Nevada, and Ohio. The result was a brief alignment of priorities, at least in terms of teacher evaluation policy, between a Democratic presidential administration and Republican state leaders. Blue states, too, passed evaluation laws. In New York, for instance, after a contentious debate with the local teachers’ union, a deal was struck to create an outcome-based evaluation system. Delaware and Connecticut also passed ambitious plans.



Source: Kathryn M. Doherty and Sandi Jacobs, “Connect the Dots: Using Evaluations of Teacher Effectiveness to Improve Policy and Practice,” NCTQ State of the States, October 2013, http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report.

CHICAGO: EVALUATION

Chicago's efforts to reform the district's evaluation system started in 2008 with the Excellence in Teaching Pilot. Prior to the pilot, Chicago Public Schools (CPS) had the same evaluation system for more than 30 years. The previous system rated 93 percent of teachers as either Superior or Excellent and only 0.3 percent Unsatisfactory, despite 66 percent of CPS students failing to meet state standards.⁸⁰

The Excellence in Teaching Pilot (EITP) better differentiated teachers based on performance than did the previous system—nearly 8 percent of teachers received an Unsatisfactory rating. The system also correlated with student growth as measured by standardized test scores.⁸¹ But in 2010, Illinois passed the Performance Evaluation Reform Act (PERA) as part of its Race to the Top bid. PERA required districts to create evaluation systems that include student growth as a "significant factor."⁸² EITP was based solely on observations, thus the district needed a new system.

CPS and the Chicago Teachers Union (CTU) eventually developed a system that met the PERA criteria: Recognizing Educators Advancing Chicago's Students, or REACH Students. REACH evaluates teachers on teacher practice, student growth, and student feedback. Trained evaluators observe teachers at least four times per year using an explicit observation rubric; observations of teacher practice count for 70-75 percent of a teacher's evaluation. Student growth, which will account for up to 30 percent of the evaluation rating, is measured through standardized tests and performance tasks, depending on the grade. And starting in 2014, student surveys will serve as the student feedback.⁸³

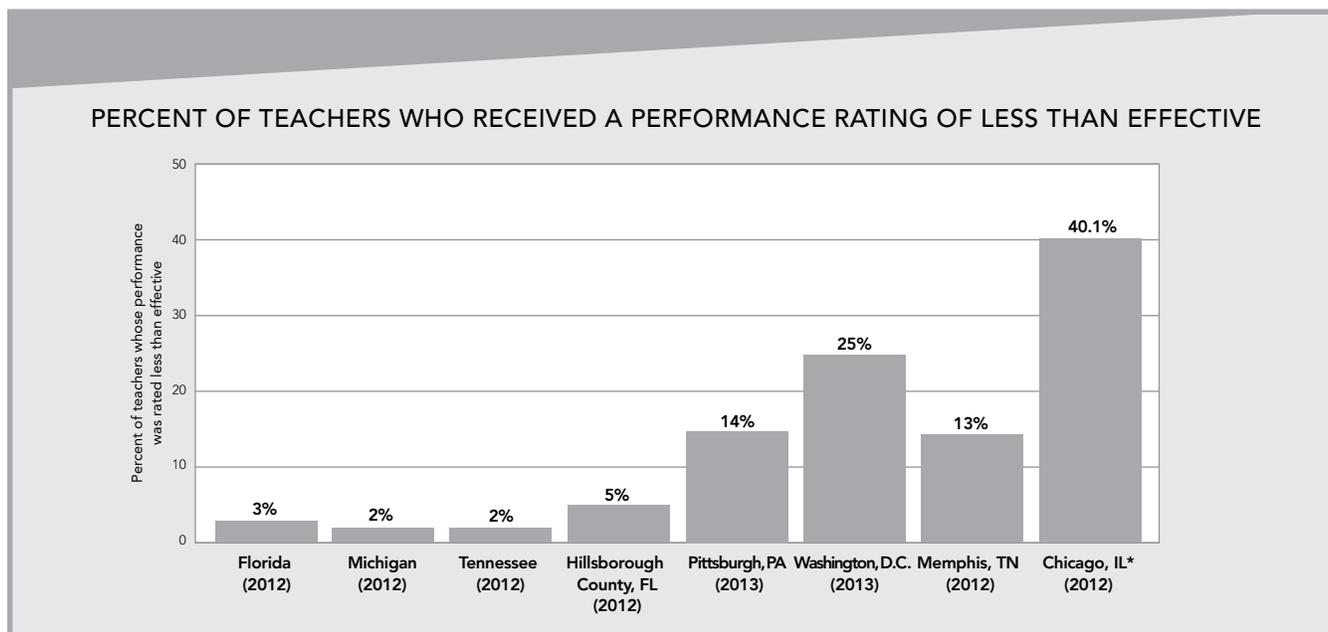
Research from the Consortium on Chicago School Research shows that REACH has been well received by teachers and principals. Seventy-six percent of teachers said the evaluation process encourages their professional growth, and 82 percent of principals reported improvement in half or more of the teachers they observed over the school year. Eighty-two percent of teachers said the new system facilitated professional conversation with their administrators, and 94 percent of principals thought the Instructional Framework improved the quality of their conversations with teachers.⁸⁴

Preliminary data also show that REACH better differentiates teacher performance. In 2009-10, before REACH, 23.5 percent of non-tenured teachers were rated Effective, 52 percent were rated Proficient, 23 percent were rated Developing, and 1.5 percent were rated Unsatisfactory. In REACH's first year, 9.6 percent of non-tenured teachers were rated Effective, 48 percent were rated Proficient, 39.5 percent were rated Developing, and 2.9 percent were rated Unsatisfactory. Compared with the other top-heavy systems, these data are promising: fewer teachers received the highest rating and more teachers received the lowest rating.⁸⁵

But in its first year, only non-tenured teachers received summative ratings under REACH. This year, REACH will be extended to tenured teachers, but under Illinois state law, tenured teachers are evaluated every two years. As a result, many tenured teachers won't receive ratings—and the full impact of REACH won't be seen—until the 2014-15 school year. Further, the district does not provide teachers with targeted, high-quality professional development to improve on their areas of challenge. A clear next step for Chicago is to focus on providing school-based, individualized coaching—based in resources from the district—to improve teaching practice.

The Department of Education subsequently encouraged similar priorities and activities in its requirements for No Child Left Behind flexibility, though the specific requirements were less prescriptive. For example, the Department only stipulated that states had to use evaluation results in “personnel decisions,”⁸⁶ not necessarily the personnel decisions that TNTP initially suggested.

States are in different stages of designing, legislating, piloting, and implementing new evaluation systems depending on their status under the multiple rounds of Race to the Top, Elementary and Secondary Education Act (ESEA) waiver applications, and their own policy choices. The result is a great deal of activity but also a great deal of variance, both in terms of specific evaluation design and ensuing consequences. Legislation in Arizona, Connecticut, Illinois, Louisiana, Maryland, Minnesota, New Jersey, Ohio, and Tennessee, for example, prohibits disclosing individual educator performance ratings. Laws in Arkansas, Florida, Indiana, and New York, on the other hand,



Source: Lisa Gartner and Cara Fitzpatrick, “Newly Released Numbers Underline Flaws in Florida’s Teacher Rating System,” *Tampa Bay Times*, December 3, 2013, <http://www.tampabay.com/news/education/k12/newly-released-numbers-underline-flaws-in-floridas-teacher-rating-system/2155404>.; Florida Department of Education, “Personnel Evaluation Data for Classroom Teachers by District,” January 18, 2013, <http://www.fdoe.org/profdev/pdf/StatewideResults.pdf>.; Venessa A. Keesler and Carla Howe, Michigan Department of Education, “Understanding Educator Evaluations in Michigan, Results from Year 1 of Implementation,” November 2012, http://www.michigan.gov/documents/mde/Educator_Effectiveness_Ratings_Policy_Brief_403184_7.pdf.; Tennessee Department of Education, “Teacher Evaluation in Tennessee: A Report on Year 1 of Implementation,” July 2012, http://www.tn.gov/education/doc/yr_1_tchr_eval_rpt.pdf.; Stephen Sawchuk, “Teacher Evaluation Sparks Clash in Pittsburgh,” *Education Week*, January 22, 2014, <http://www.edweek.org/ew/articles/2014/01/22/18pittsburgh.h33.html>.; Data from the Office of Human Capital, District of Columbia Public Schools.; Lisa Watts, Joy Singleton-Stevens, and Mark Teoh, “Lessons from the Leading Edge: Teachers’ View on the Impact of Evaluation Reform,” 2013, http://www.teachplus.org/uploads/Documents/1371738773_lessons_from_the_leading_edge.pdf.; Rebecca Harris, “New teacher evaluations get positive reviews,” *Catalyst Chicago*, September 18, 2013, <http://www.catalyst-chicago.org/notebook/2013/09/18/60249/new-teacher-evaluations-get-positive-reviews>.

*Chicago data refer only to untenured teachers.

explicitly require aggregated educator performance data. Legislation in Arkansas, Connecticut, Delaware, Maryland, Minnesota, New Jersey, and New York does not require teacher performance to be the primary factor in layoff decisions, and only Colorado requires mutual consent in placement of excess teachers.⁸⁷

As new evaluation systems are implemented, the evidence is mixed about whether they are making more meaningful distinctions based on performance than the systems they are replacing. In some places, teachers seem to be getting the same ratings. For instance, in Hillsborough County, Fla., only 1 percent of teachers were rated Unsatisfactory and more than 95 percent of teachers were rated Effective or Highly Effective in 2012.⁸⁸ Yet Hillsborough was expected to be an example of success. It is frequently lauded as a leader on evaluations and management-labor collaboration and received a \$100 million grant from the Bill & Melinda Gates Foundation to support its teacher quality work.⁸⁹ Statewide, nearly 97 percent of teachers in Florida were rated effective or highly effective under the first year of the new teacher evaluation system.⁹⁰ Michigan⁹¹ and Tennessee⁹² each saw 98 percent of teachers rated effective or better during the first year of those evaluation systems.

But in others, the distribution of ratings looks quite different. In Pittsburgh, Pa., by contrast, the school system's new evaluation system—developed jointly with the teachers' union and also supported by the Gates Foundation—identified 14 percent of teachers in the lowest two categories for performance. The teachers' union objected to the new measure, saying that figure was 10 times the national average.⁹³ In 2013, 25 percent of teachers in Washington, D.C., received a rating below Effective on the IMPACT evaluation system.⁹⁴ In 2012, 13 percent of Memphis teachers⁹⁵ and 40.1 percent of Chicago untenured teachers⁹⁶ were rated in the bottom two categories of effectiveness.

Other evidence on teacher evaluation suggests new systems may lead to increased separations of low-performing teachers. Data from Washington, D.C.'s IMPACT show that 418 teachers have been dismissed because of low performance ratings since 2009.⁹⁷ And data from Chicago and Tennessee indicate that teachers perceive evaluation systems positively. In 2013, 53 percent of Tennessee educators agreed or strongly agreed with the statement, "In general, I believe the teacher evaluation used at my school will improve teaching." Only 38 percent of educators in Tennessee agreed or strongly agreed with that statement in 2012. In Chicago, 76 percent of teachers said the new evaluation system encourages professional growth. Eighty-eight percent of teachers said their evaluator was able to assess their instruction accurately, and 93 percent of administrators said the Chicago framework is useful for identifying teacher effectiveness.⁹⁸

The existing data on student achievement outcomes are limited but promising. The 2013 NAEP scores in Washington, D.C., and Tennessee show that these areas had the largest gains nationally since 2011.⁹⁹ The causes are unclear and may not be linked at all to teacher quality reforms—D.C.’s scores, for example, include the city’s large charter sector. But it’s worth noting that two of the earliest adopters of reformed evaluation practices now lead the country in NAEP growth.

Given the weak condition of professional development, states and school districts are even less successful in targeting support or growth opportunities to teachers under new evaluation systems. The problem is not a lack of models. Research organizations, think tanks, and the teachers’ unions have proposed strategies to link evaluation with professional development. Nor is it the obviousness of the need to ensure that support for teachers is differentiated and focused based on evaluation results. The challenge in most states is the simultaneous effort to implement new evaluation systems, implement new standards with the Common Core or other career and college-ready standards in states not adopting the Common Core, and the existing lack of capacity for professional development.

PREPARING AND COMPENSATING TEACHERS: WHERE WE ARE NOW

Education is in the midst of significant changes to its approach to human capital and must also respond to evolutions in the larger labor market. However, traditional teacher preparation programs still look much the same as they did several decades ago: candidates complete required coursework—including foundational, pedagogical, and subject matter content—and field experiences. Alternative routes may get the headlines, but the majority of prospective teachers enter the profession through traditional programs, often housed in schools of education.¹⁰⁰ As of 2013, traditional programs produced 200,000 graduates every year.¹⁰¹ In some states they produce more teachers in certain subjects and grade levels than there are jobs available.

There is wide variation in the quality and composition of traditional teacher preparation programs, depending on the state; programs are shaped by state regulations, accrediting bodies, and institutional and program choices. But there is a fairly consistent belief that the quality of traditional preparation programs is lacking. Critics disagree on remedies but there are critics within and outside of the academy and across a wide spectrum of education leaders.

Once trusted sources of expertise, teacher-training institutions are now increasingly questioned. “The credibility of universities to be the arbiters of teacher preparation is at an all-time low,” TNTP’s Tim Daly says. “It has gotten to the point where everything they say is assumed to be untrue because they are perceived as feckless ideological defendants.”¹⁰²

Critics of university-based teacher preparation argue that states pay little attention to quality control, laborious accreditation practices do not affect teacher quality, and program completers are still unprepared. Programs often ignore labor market issues, producing too many elementary school teachers and too few STEM and special education teachers or those willing to teach in high-need areas.¹⁰³ Surveys of graduates from many of these programs indicate they feel unprepared when they arrive in the classroom, and that their training was insufficiently rigorous. And research shows that there is not an appreciable difference in student achievement between teachers with traditional preparation credentials and those without.¹⁰⁴ Except for emergency credentials, research shows candidate characteristics matter more than program characteristics to classroom effectiveness. In other words, this highly regulated approach is not even generating better outcomes than more efficient alternatives.

Recent research—the first of its kind—provides data to support the anecdotal conclusions of preparation program critics. The National Council on Teacher Quality (NCTQ) developed program standards of quality in 10 pilot studies over eight years. In 2013, NCTQ published a report evaluating 2,420 teacher preparation programs in 1,130 institutions on these standards. Out of four possible stars of quality, the paper rated fewer than 10 percent of programs at three stars or above. Some of the issues identified were easy admission standards, inadequate content knowledge, poor classroom management skills, and ineffective clinical partner teachers.¹⁰⁵

Little consensus about the best approach for reforming traditional preparation programs exists. Few efforts exist, and most are in their nascent stages. For example, the Council for the Accreditation of Educator Preparation (CAEP) is in the process of increasing its accreditation standards, and some states promised to link program approval to student outcomes as part of their Race to the Top application. It's too soon to know how consequential these steps will be.

Meanwhile, alternative route programs, like Teach For America and TNTP, launched and grew. Nationally, 31 percent of all teacher preparation pathways are alternative route programs.¹⁰⁶ Forty-seven states allow alternative routes to teacher certification.¹⁰⁷ Although they vary in quality, these programs generally offer condensed, “boot camp”-style training for participants before they enter the classroom, followed by regular professional development throughout the program duration. The debate is fairly contentious on alternative route program quality, but is based in ideology rather than research. Multiple studies show that high-quality alternative route teachers, particularly those trained by TFA, perform as least as well as other teachers, including veteran teachers, and, in some cases, slightly better.¹⁰⁸ These outcomes raise questions about the costs and benefits of traditional teacher prep programs as well as broader questions about how to improve the preparation of teachers overall.

COMPENSATING TEACHERS BASED ON EFFECTIVENESS

Analysts disagree about whether teachers are appropriately compensated or underpaid. The debate largely turns on different points of comparison and what aspects of teacher compensation (for instance, pensions) are included as a measure of compensation.¹⁰⁹ Compensation also varies tremendously by geography, so average teacher pay is often not a useful statistic. However, a consensus has emerged that, based on their impact on student learning, the best teachers are generally underpaid and under-recognized.

While the debate over teacher pay is heated, it is largely disconnected from compensation policies in the vast majority of school districts. Overwhelmingly, teacher compensation remains based on a district-wide single-salary schedule that rewards years of experience and academic degrees. Despite ample evidence that years of experience and advanced degrees have little or no bearing on student achievement, those are more often than not the factors that dominate teacher compensation.¹¹⁰

Merit pay, performance pay, or other terms for rewarding teachers based on their performance in the classroom are not a new idea. Despite the certainty of critics that the idea does not work and the certainty of many of its proponents that it will boost achievement, the research to date is inconclusive in no small part because robust initiatives to pilot the idea are so rare. The best-known example is the system that the District of Columbia Public Schools included in its 2010 teacher contract. DCPS's system, *IMPACTplus*, allows teachers rated as Highly Effective on the district's evaluation system—which heavily weights student growth—to opt into the performance-based compensation structure.

Under *IMPACTplus*, Highly Effective teachers can earn nearly double what they would have earned in their first year and can achieve one and a half times the previous maximum salary in less than half the time. Teachers who accept bonuses under *IMPACTplus*, however, are ineligible for the buyout or “extra year” provisions granted to other teachers if they are later exceeded.¹¹¹ Preliminary research on *IMPACTplus* suggests that the financial incentives may improve the performance of high-performing teachers.¹¹² Denver's ProComp system, introduced in 2004, follows a similar structure with smaller financial incentives and indirect links to student achievement. Preliminary research from ProComp is also positive: student outcomes, teacher retention, and teacher recruitment have all increased, and nearly 75 percent of teachers participate in the program.¹¹³ Evidence from a performance-pay pilot in Arkansas found positive effects on the lowest-performing teachers.¹¹⁴

A Vanderbilt School of Education analysis on a differentiated compensation system in Nashville found no lasting effects on student achievement but also no adverse impact on teachers or students.¹¹⁵ Douglas County School District, also in Colorado, is experimenting with a market-based compensation structure, in which teachers with credentials in high-need or undersupplied areas are compensated higher than teachers with credentials in oversupplied areas. There is no research to date on the effectiveness of Douglas County's model.

A decade ago, the idea of differentiated pay, even based on especially challenging assignments, scarcity, or other factors besides student outcomes, was highly controversial. Ambitious pilots remain rare and as a result it's still unclear exactly which design elements are most or least effective in a differentiated pay scheme. The idea that salary should be more differentiated based on measures of contribution is slowly gaining acceptance within teachers' unions as well as the policy community. Performance-based rewards or incentives may not spur higher levels of performance but do send important signals about how the profession considers and rewards performance. There is no reason to believe that the incentives that matter in other areas of American life do not also matter, at least to some extent, within education. As measures of teacher effectiveness and differentiation based on effectiveness become more embedded, money seems likely to begin to follow those measures, too.

PENSIONS

Collectively, using the states' own figures, the gap between what states have saved for teacher pensions and what they have promised totals \$390 billion. In some communities these shortfalls are beginning to pressure current budgets. For instance, Mayor Chuck Reed of San Jose, Calif., is vocally campaigning for pension reform because of the impact of the retirement system on his city's budget. States have responded to those shortfalls by increasing district costs and enacting punitive policies that reduce the amount teachers can expect to receive from a pension and make it harder for teachers to qualify for a pension at all. State and local governments have made their pension systems less friendly to young and mobile workers by lengthening vesting periods and creating separate, less generous plans for new employees.

While the basic structure of teacher pension systems remains largely the same as when they were first adopted, the teaching profession itself has changed a good deal over the past quarter century. Where teaching was once a relatively stable profession and retirement systems that favored longevity suited a large portion of the workforce, today there's significant mobility among teachers and few will reap the benefits of the pension system. According to state figures, half of all Americans who teach in public schools won't qualify for even a minimal pension benefit, and fewer than one in five will stay long enough to earn a normal retirement benefit. Teachers who leave the profession or who move across state lines face significant savings penalties. Those penalties can amount to a few thousand dollars if they leave after one year or hundreds of thousands of dollars if they split a 30-year career in two or more pension systems.

This is not a marginal problem. Numbering 3.3 million, public school teachers constitute the largest class of college-educated workers in the country. In other words, policymakers are systematically disadvantaging our largest class of bachelor's-degree-equipped workers as well as making the profession less attractive as a career choice.

WHAT'S NEXT? RECOMMENDATIONS

“Previously, as a superintendent you could focus on change without focusing on teacher quality. Now you have no other choice.”

- Talia Milgrom-Elcott, a former program officer, Carnegie Corporation for New York

America is finally having a hard conversation about its teachers. The past decade saw an enormous shift in how policymakers and philanthropists consider teachers and led to changes that, if successfully enacted, have the promise to move K-12 teaching toward a genuine professional approach and make the work of teachers more respected and satisfying. “Previously, as a superintendent, you could focus on change without focusing on teacher quality,” said Talia Milgrom-Elcott, a former program officer with the Carnegie Corporation of New York. “Now you have no other choice.”¹¹⁶ Yet despite the widespread recognition of the need for change, the pace of action is slow, and for every high-profile model there are literally hundreds of school districts

whose practices remain largely unchanged. Even where there is change, variance in the systems rather than consistency defines the status quo. It’s impossible to accurately speak of “teacher evaluation” in the singular. States are trying a variety of different strategies with broad and subtle differences amongst them.

It is clear that the next decade requires a more talent-minded and research-based vision for American teachers. What is unclear is the pace. Some leaders within the field worry that teacher quality policy has moved too fast, outpacing implementation, but Secretary of Education Arne Duncan disagrees. “Teacher quality policy has moved,” Duncan says. “But it has moved far too slowly.”¹¹⁷

The slow progress stems in no small part from the politics and battles necessary to make changes. The 2007 New York City¹¹⁸ and 2010 Washington, D.C.,¹¹⁹ teacher contracts ended forced placement and required mutual consent hiring, but the New York City contract left the district with an annual \$100 million obligation to pay nonworking teachers in Temporary Reassignment Centers.¹²⁰ Both contract negotiations were nasty, brutish, and anything but short, creating a disincentive for others to take on these fights. In Chicago, the city’s teachers went on strike for eight days over a variety of issues, including evaluation and job protections.

Some, though not all, of this is the inevitable nature of change in the education sector. Much of the teacher quality conversation of the past several years has focused on teacher evaluation, particularly evaluating teachers after they start teaching and, essentially, dismissing them if they are not effective. There is undoubtedly too little attention to underperformers and their deleterious

effect on student learning and school culture, yet the dismissal of teachers is not the end goal of policy or a sufficient reform. The teachers' unions have an interest in stoking a climate of anxiety among teachers in order to position the union as a needed resource for them. But they have too often had a partner in education reformers who have failed to put forward a holistic narrative that is as much about celebrating and supporting exceptional teachers (who far outnumber low-performers) as it is about forcing hard decisions about underperforming teachers. The conversation must be as much about lifting the ceiling as it is about bringing up the floor.

In addition to fostering to a balanced narrative, reformers should continue to encourage as many teacher voice organizations as possible to broaden the debate, regardless of the specific positions these groups take on particular policy questions. Recent survey research indicates that teacher voice is substantially less monolithic, and that teachers' unions represent the views of fewer teachers, than in previous years. Teacher voice organizations follow various models—Teach Plus, America Achieves, and Hope Street Group have cohort fellowships, while Educators 4 Excellence and VIVA build networks of engaged educators and NNSTOY focuses on state teachers of the year—and add an important diversity of viewpoint.

Within that broader conversation, here are five areas requiring careful attention from policymakers, analysts, philanthropists, and advocates:

You can't people-proof systems in education

The past century of education reform is littered with efforts to take human judgment out of educational decisions and practice. These strategies try to “teacher proof” reforms or address politics and the inability or unwillingness of educational managers to make tough decisions. It rarely works—especially given education's decentralized nature, which offers plenty of opportunities to evade objectionable requirements and fundamentally relies on people.

The prevailing approach to teacher evaluation today, while a vast improvement over previous policies, echoes the idea of taking people out of the equation. New evaluation systems generally rely on professional judgment but are still light on professional discretion. This may explain why in some communities new evaluation methods are producing the same inflated results as the policies they replaced.

Evaluation efforts are developing elaborate metrics in an attempt to force tough decisions. These tools are valuable and can provoke real conversations about practice and effectiveness in the classroom. But while they are ostensibly based on professional judgment, they can leave little room for genuine professional nuance and respect for the grey areas common in professional work.

This approach is an understandable response to the “widget effect” problem and arguably an unavoidable step toward the professionalization of education. In addition, because teachers’ unions resist approaches allowing a great deal of discretion, fearing the possibility of favoritism and other abuses, it’s an unsurprising place for the field to find itself in. While teachers’ unions and reformers frequently disagree about the design of evaluation systems, there is nonetheless some consensus around the metric-driven approach. Reformers like metrics because they trust in them as a way to force action and unions favor them because they do not trust managerial discretion and insist on a codified and rules-based framework for managers. It perpetuates a compliance-based ethos and reinforces an adversarial relationship between management and labor.

In most kinds of professional work, evaluation is a function of an evidence base, frequently built with formalized tools and processes, coupled with managerial discretion and corresponding accountability for high-quality decision-making. In other words, managers consider the data, use their judgment, and are held accountable for the aggregate quality of their decisions overall. In education, this approach exists in some charter schools and private schools, but it’s scarce within the traditional public system. Somewhat ironically, for all the decrying of “corporate” education reform, it is the private sector that offers some of the best practices for professional, context-specific, respectful, and meaningful evaluations—that don’t rely on standardized tests.

Unions complain that this type of evaluation and personnel management can lead to teachers losing their jobs for their political views, sexual orientation, or religion. But while there is always a risk that power will be abused, a host of federal and state laws protect most teachers from discrimination in the workplace. The larger problem facing education is not abusive management, it is management that has never been trained to conduct and use evaluations in a performance-driven and rigorous way, or be held accountable for doing so. School managers are beginning to get that training now, as a result of the evaluation thrust, but compliance with a new set of rubrics will not build a genuinely professional culture.

Public education must move past its attachment to completely metric-driven evaluation approaches and embrace the messy norms that occur when managerial discretion (and the concurrent accountability) is not only allowed but actively encouraged. This is, of course, at odds with the language in many teachers’ contracts and counter to much of the culture in education—especially within the unions. However, such an approach would professionalize the principalship (still a relatively low-autonomy job) and the teaching field, and it would bring education more in line with other professions’ evaluation strategies.

Managerial discretion does not mean jettisoning today’s rubrics, tools, or the appropriate use of value-added data or other measures. Nor does it mean anything goes. Instead, it means allowing principals to use these data to inform their evaluative decisions while factoring in all the qualitative

and professional judgment that comprises a true professional evaluation. Arguing over whether value-added scores should comprise 35 percent or 51 percent of a teacher's evaluation is less important than developing a system that actually holds principals accountable for removing low-performers, growing high-performers, and gives them the autonomy to do so. Considering that most teachers do not even teach in subjects or grades assessed by standardized tests, it's remarkable how much value-added has dominated the conversation about evaluation at the expense of a broader conversation about how professionals evaluate one another and improve their work.

Jal Mehta and Steven Teles have suggested an approach of "plural professionalism," allowing different networks of schools to evaluate teachers based on their shared values. Plural professionalism suggests there is not just one uniform professionalism that fits all schools. Instead, schools that share similar values fit into a network that also shares the components of professionalism necessary to support those values: similar approaches to training, curriculum, assessments, and accountability.¹²¹

For policymakers, encouraging managerial discretion means allowing carefully crafted waivers from today's evaluation schemes in the near term for clusters of schools, whole districts, or consortia. In the long term, it means revising these systems as new models emerge and constraints on managerial discretion ease. For philanthropists, it means investing in pilots to train and support principals in best practices for evaluation as well as larger investments in improving school leadership. For reformers, it means being willing to walk back some of today's rigid systems in favor of more professional and collaborative models. Teachers' union leaders must be willing to exchange a system based on rules and compliance for more professional norms, heightened accountability for outcomes, and honest conversations about effectiveness, and to lead their members toward a more professional place. If lawsuits such as *Vergara v. California* succeed in striking down common and outmoded provisions in state law or teachers' contracts, the unions will be forced to take steps in this direction. Yet the education system is arguably not ready for such approaches at any scale.

Professionalize professional development

Teachers are frustrated that most evaluation systems are not well-aligned with meaningful opportunities for professional growth. Evaluations should not merely be about addressing low performance. They must be linked to incentives for success and systems to spur professional growth and learning.

Lost in the din about the teacher evaluation system in Washington, D.C., for instance, is the high degree of support from teachers in no small part because of the opportunities IMPACT*plus* offers for substantially higher pay and increasingly for professional growth.

Teachers are understandably exasperated by ineffective professional development. Data from the 2013 *Primary Sources* survey, a national survey of more than 20,000 pre-K through 12th-grade teachers, show that only 60 percent of math and English teachers and 46 percent of science and social studies teachers found professional development useful or very useful.¹²²

The large educational publishers are a frequent target of criticism for the low quality of professional development in education. If they were the only culprits, it would be an easier problem to address. In practice, across the education sector a diverse range of providers, states, and school districts conduct professional development of varying quality. Many of the providers are small, frequently former teachers or school officials themselves, and quality control is largely nonexistent. The problems range from a lack of customization to a lack of quality and rigor.

Federal policymakers have sought to improve the use of professional development dollars but the variety of programs, lack of coordination amongst them, and decentralized nature of education have thwarted efforts to substantially improve professional development. In FY13, more than a dozen federal programs funded professional development activities. The largest federal program dedicated to K-12 teachers—the Improving Teacher Quality State Grant program—allocated nearly 45 percent of its \$2.3 billion to professional development.¹²³ But the Improving Teacher Quality State Grant program allocates funds to state educational agencies, which then must pass the funds through to districts. Each district pursues a different professional development strategy, producing an unwieldy range of programs. Similarly, between 2010 and 2012, the Investing in Innovation (i3) program awarded \$937 million in federal competitive grants. More than half of those funds—\$457 million—went to 62 different projects that relied on professional development as a key lever.¹²⁴

For policymakers, educators, and philanthropists, the professional development challenge is threefold and a balancing act of what’s actionable today and what’s needed for tomorrow.

Fifty-seven percent of teachers said they received less than 16 hours, or two days, of professional development in their content area in the past year.

First, the field must rigorously identify and develop models of professional development that improve teacher effectiveness and student achievement. The existing literature on effective models of professional development is sparse, but research suggests that teachers benefit most when they receive intensive, ongoing training that connects to a specific discipline or grade level.¹²⁵

A 2007 review of nine evaluations of K-12 professional development showed that a substantial investment of time in professional development positively affected student achievement. The average professional development duration (49 hours) boosted student achievement by 21 percentile points. Studies that examined programs with less than 14 hours of

professional development found no statistically significant impact on student achievement.¹²⁶ And yet, most teachers receive short, episodic professional development. Fifty-seven percent of teachers said in 2004 that they had received less than 16 hours, or two days, of professional development in their content area over the past year.¹²⁷

Other research, however, suggests that duration and specificity do not guarantee effective professional development. In 2011, the Institute of Education Sciences (IES) released the results from a two-year professional development program designed to help 7th-grade math teachers teach rational numbers. The evaluation showed that professional development did not have a statistically significant impact on either teacher knowledge or student achievement.¹²⁸ Another two-year model focused on early reading also showed that, by the end of the second year, the program did not have a statistically significant impact on teacher practice or student outcomes.¹²⁹

The “lesson study” method, adapted from Japanese professional development, is becoming more popular. In a 2011 evaluation, 213 teachers split into small groups. Each lesson study group met 12-14 times over five months to collaboratively develop, teach, and analyze fractions lessons for 1,059 students. At the end of the evaluation, both teachers and students showed statistically significant increases in knowledge of fractions.¹³⁰ Rather than a rigid adherence to the Japanese lesson study, a model to emulate the elements of this approach—sustained, tailored, and relevant—is key. In almost any profession successful training and development can involve different approaches if they are rigorous and aligned with the underlying work. That sort of work, however, has to come from the bottom; it cannot be mandated.

Seventy-six percent of Chicago teachers said the district’s new evaluation system encourages their professional growth.

Second, the field needs evaluation systems that align with professional development and emphasize teacher improvement as well as performance management. Not surprisingly, teachers view evaluation systems that improve their professional practice more favorably. In a 2013 study, 76 percent of Chicago teachers said the new evaluation system encourages their professional growth.¹³¹ According to the 2013 *Primary Sources* survey, teachers are more likely to find their school’s evaluation system extremely or very helpful if they receive customized professional development after an evaluation—but only 13 percent of teachers do.¹³² Building these linkages is critical and should be a point of common ground.

More controversial are proposals to couple investments in professional development of proven teachers with a focus on performance management for novices. This approach conflicts with the conventional wisdom that struggling teachers will eventually improve into effective teachers. Yet evidence supports innovation. Research suggests that teachers improve during their first few years of practice,¹³³ but peak after three to five years.¹³⁴ In a recent analysis, TNTP showed that first-year teachers were, on average, more effective than third-year low-performers, despite the difference in experience.¹³⁵ Educators and policymakers should innovate with when and how to invest professional development dollars rather than spreading them around too thinly to have impact. The data indicate that investing in teachers at the lowest-performance levels rather than removing them may not be cost-effective and may dilute scarce resources that would be better spent on other teachers.

Finally, improving professional development for tomorrow is a challenge of innovation more than policy. Improved online education offers the chance to break down traditional time and place barriers. The best professional development can be delivered virtually with higher fidelity and at lower costs. At the more leading edge, there are early-stage efforts to use virtual reality to help teachers practice and hone their craft. Similar to what pilots experience, teachers could practice their skills in simulations instead of real-world situations. The Common Core State Standards, adopted by 45 states, can facilitate greater scale in professional development because teachers will work to solve similar problems, with similar standards and curricular materials, across various school districts and states. There are early-stage ventures attempting to develop new approaches and the Bill & Melinda Gates Foundation is in the early stages of seeking new ideas and approaches. It's a vast greenfield for innovation.

Open and expand teacher preparation

Reforming teacher preparation is instrumental to improving overall teacher quality. The past decade's efforts focused primarily on alternative routes, essentially attempting to bypass traditional teacher preparation programs. These efforts have made a difference and highlighted important issues in teacher preparation, but are insufficient to the scale of the teacher labor force. With so many prospective teachers completing traditional programs, the heavy focus on alternatives in the political debate ignores the majority of teachers entering the classroom through traditional routes. "We are very hesitant to take on the remaking of teacher preparation," Kati Haycock, president of the Education Trust, says. "We've been adding Band-Aids, like TFA, but not addressing the fundamental problem."¹³⁶

That's because teacher preparation remains such a difficult part of the sector to reform. Structural challenges limit the leverage policymakers have, economic considerations dissuade universities from addressing problems, and the politics of teacher preparation are as thorny as any in the sector. Eric Hanushek, senior fellow at the Hoover Institution, observes, "Education schools also do not have a culture of evaluation where they assess how well their programs are doing and how any changes relate to student effectiveness."¹³⁷ Jack Jennings, founder of the Center on Education Policy, says preparation programs are often "cash cows." He notes, "The university makes money off of them without increasing the quality of the teaching force."¹³⁸ Yet a labor-intensive field like education is only as strong as the people going into it. Policymakers must:

Increase rigor and choice in teacher preparation. "It's a problem that teachers are coming from the bottom third academically," Secretary of Education Arne Duncan said. "That hasn't moved much at all in the past decade but it needs to move radically."¹³⁹ As a rule, educationally higher-performing countries draw a more selective cohort into teaching. And preparation programs are the gatekeepers to the world of teaching. Programs that admit and graduate candidates with low academic achievement and who lack subject-area content knowledge perpetuate the problem of ineffective teachers—both in perception and in reality. But low standards carry a cost for some teaching candidates as well. In states with more demanding standards, these candidates can complete a preparation program but then fail to pass gateway exams.

Higher admission standards, subject-area major requirements, and restricted use of admission waivers can help increase the minimum level of candidate performance before they even enter their first teacher preparation class.

Respect the evidence and end protectionism of traditional preparation programs. There is evidence that clinical fieldwork can benefit prospective teachers. The research on the ideal amount and design of clinical practice is inconclusive, but the National Council on Teacher Quality (NCTQ) suggests that candidates should experience at least 10 weeks of clinical practice in a classroom with an effective mentor teacher who has been teaching for at least three years.¹⁴⁰ Other research offers some support for the idea that programs focusing on actual classroom work produce better outcomes.¹⁴¹ But there is also evidence suggesting that candidate characteristics matter more than the particular route aspiring teachers follow.¹⁴² And today there are still more questions than answers about effective teacher preparation despite elaborate regulatory and credentialing regimes and millions of dollars spent annually on these programs.

States should release outcome information about different preparation programs to ensure heightened transparency. At the same time, states should allow for greater choice among programs by teachers and by schools. Coupled with transparency, allowing schools to hire teachers from a wider range of training programs and allowing candidates to choose from a greater range

of options will add additional pressure for all programs to improve, create more options and innovation for teacher preparation, and increase pressure specifically on low-performing programs. Alternative routes that meet a threshold for rigor, which existing routes such as TNTP, Teach For America, and various state and local alternative paths do, should be able to compete for students alongside traditional programs as they do in some states now. More pluralism in credentialing and the resulting change in the human capital profile has played a role in the rapid improvement of schools in cities such as Washington, D.C., and New Orleans.

“Colleges of education should be shut down if they aren’t producing effective teachers,” says Jennings.¹⁴³ Doing that through regulations has proven largely ineffective. The dual pressure of public transparency and greater choice for teaching offers more promise.

Address productivity

Engaging with productivity-enhancing measures seems obvious. In practice, owing to politics and tradition, reform in American education has been largely additive and avowedly at odds with improved productivity for more than a quarter century. New initiatives and new people are layered on top of existing arrangements. Hard decisions are generally only made in times of genuine scarcity.

Nowhere is this truer than for teachers. For example, research on class-size reduction shows that it most benefits students in the early grades and high-need students. Yet policies reducing class size have spread dollars across school districts. All classes are reduced by a student or two rather than through targeted reductions for the students who most need it, resulting in negligible impact across the board rather than meaningful impact where the policy could be most effective. Focusing on quality rather than quantity would also leave teachers better paid for their efforts than they are today.

What might a productivity-focused agenda look like? The goal might be to capitalize on the existing pool of teacher talent. Obvious steps include ending seniority-based layoffs and encouraging schools to retain their best performers in the classroom when layoffs occur. In the overwhelming majority of school districts, seniority—not performance—determines layoff decisions. NCTQ studied 100 large districts and found that in all but 25 of those districts, seniority is the primary determinant in teacher layoff decisions.¹⁴⁴ But that trend is beginning to change: as of 2012, some 10 states require teacher effectiveness to inform layoff decisions, and three states prohibit layoff policies based on seniority alone.¹⁴⁵

Policies could also focus on recruitment. Pension wealth could be spread more evenly across a teacher’s career to raise salaries in the earlier years rather than back-loading pension wealth until the last few years. Or late-career salary dollars could be shifted to substantially raise the salaries of early-career teachers, perhaps to make the “tenure” point a significant bump, as it traditionally was in higher education.

At a minimum, schools should pursue “hire slowly and fire quickly” human resources strategies with novice teachers. Research indicates that schools that dismiss a low-performing teacher have a 73 percent chance of replacing them with a more effective teacher. Yet they routinely fail to do so. Each year, nearly 10,000 of the best-performing teachers—those who help students learn between 2-6 months more than other teachers—leave the 50 largest school districts. At the same time, nearly 100,000 low-performing teachers stay.¹⁴⁶ The result saddles schools with ineffective personnel who adversely impact students and sap the morale of millions of good teachers.

Technology can also help. Public schools remain one of the last large American institutions largely untouched by productivity-enhancing reforms from new technologies. Computers, laptops, tablets, and video will not, and should not, replace high-quality live teaching. But there are places where technology can free teachers to focus their work better, support them in the classroom, or reinforce basic skills so teachers can tailor their efforts elsewhere. The benefits of “ed tech” are arguably oversold, but technology does have a key role to play in schooling. Already, new models help teachers and schools operate more productively and effectively. For example, the “New Classrooms” and “Rocketship Education”¹⁴⁷ approaches to teaching allow teachers to customize instruction across many students, and other online programs and interventions show promising results.¹⁴⁸ Public Impact, a North Carolina-based education consulting firm, is developing an “opportunity culture” approach to finding ways to extend the reach of the best teachers through different tools and modalities.

Address the politics

The politics of education are challenging, particularly when it comes to human capital and teacher effectiveness. Teachers’ unions and associations are potent political forces at the local, state, and national level. They are not only the largest political force within education, but they are also among the very biggest spenders on political activity in the country.¹⁴⁹

Two basic dynamics underlie the politics of the education debate:

- In American politics, it’s easier to block change than to create it.
- Special interests tend to trump the general interest because they are focused, there every day to advocate, and organized.

Although the internal dynamics of the unions make it difficult for them to embrace reform, it can happen. Yet these instances are rare because of basic organizational dynamics. Like most organizations, the most strident members drive teachers' union elections. Union elections are also driven by the politics of the present, not the possible politics of the future. In other words, the politics within the unions are driven by the membership of today (who vote), not the future interests of tomorrow (that don't). Union leaders may see a threat down the road, but convincing members to respond to that threat, especially when the response carries immediate costs, is a tough sell. It's no different than the issues politicians face trying to get voters to bear costs to address global warming or reform entitlement programs. The present costs are clear, while the benefit is only apparent at a future time—always a tough sell in politics.

For the teachers' unions there is no issue more front and center than the jobs of their members. Teachers' union leaders are not inherently uncaring about other concerns, but the organizations they lead are, at their core, membership organizations that elect their leadership. Union leaders are politicians just as much as the elected officials they lobby.

Where substantial reform has happened, however, two lessons can be drawn.

Circumstances, not desire, dictate reform. In every case where unions have embraced reform, external context has helped force the conversation. In New Haven, for instance, the mayor was prepared to act unilaterally. For the teachers' union, the choice was to take a seat at the table or be marginalized. In Washington, D.C., teachers wanted to embrace the new contract and the salary dollars that came with it, so the union was forced to embrace a contract that contained reformist language on evaluations. Conversely, in Chicago, the teachers' union felt no pressure to compromise, went on strike, and empowered hard-liners in teachers' unions around the country. When reform is happening, people are at the table so they can have a say. When it's not, they're not.

There are no national models, only local politics. When New Haven adopted its new contract, union leaders hailed it as a national model. American Federation of Teachers President Randi Weingarten called it "the gold standard."¹⁵⁰ Yet New Haven has not been replicated elsewhere. On the contrary, during high-profile debates about evaluation in cities across the country, approaches similar to New Haven's were not on the table. This speaks to the localized nature of many of these debates. Just because Weingarten or any other national leader says a policy is the ideal does not mean a local union leader elsewhere must embrace it. And teachers' union leaders are elected locally, so they're much more attuned to their local constituents' desires than are national leaders, pundits, analysts, or advocates.

Politics matter because at its core the political debate about human capital issues has little to do with empiricism and a great deal to do with, well, politics. The substance matters to policy design, but the political debates do not turn on the force of analysis. The gap between the rhetoric about value-added data and the actual way states and school districts use those data illustrates the political nature of this debate, as do the never-ceasing calls to mute the effect of these evaluation systems. Absent an organizing and political strategy, supporters of revamping teacher quality policy will always be playing defense. The familiar educational pattern of new idea, followed by weak implementation in the face of political pressure, and subsequent discrediting of the idea will persist.

A successful political environment favors and supports innovation rather than seeking to thwart it. Leaders need political breathing room to operate. New ideas need political support to not only launch but be sustained long enough to see if they work. Innovators, social entrepreneurs, and policy entrepreneurs need political space to operate. Initiatives need political support so that all the sharp edges are not sanded off through the political process. And most importantly, ideas need time to be developed and the political zigging and zagging of teachers' union leaders and many elected officials is at odds with any innovation cycle.

To date, the education reform community has largely tried to change education politics without actually doing the fine-grained work of politics. Candidate-based efforts, backed by organizations like Democrats for Education Reform, helped reform-sympathetic politicians win office. But on the ground, in terms of grassroots politics, organizing, and localized action, the teachers' unions remain the dominant force in local politics. The political battle largely ends up as white papers and breakout sessions versus skilled political organizing. As these local political fights erupt, the teachers' unions deploy political operatives with deep experience in campaigns—they know how to do politics. There are exceptions—in New York City, charter school leader Eva Moskowitz has proven adept at organizing parents to advocate for their children at city council hearings and in the state capital. Some education advocacy organizations at the state level are having success as well. But in general all other actors in education play for second place to the political operation of the teachers' unions. Reformers need not match the unions dollar for dollar but must change the organizing imbalance if they hope to genuinely change urban education.

Philanthropists cannot directly support some kinds of political advocacy with tax-exempt philanthropic dollars, but must ensure that there is a political strategy supporting the reforms they favor. To do less is similar to buying a house but failing to buy homeowners insurance to protect it. Until reformers level the playing field and change the fundamental special interest versus general interest dynamic that characterizes education politics they can expect to play defense or at best see incremental progress—especially on human capital issues.

ENDNOTES

- ¹ Valerie Strauss and Mark Naison, “The War on Teachers: Why the Public Is Watching It Happen,” *The Washington Post*, March 12, 2012, http://www.washingtonpost.com/blogs/answer-sheet/post/the-war-on-teachers-why-the-public-is-watching-it-happen/2012/03/11/gIQAD3XH6R_blog.html.
- ² Ann Duffett et al., “Waiting to Be Won Over: Teachers Speak on the Profession, Unions, and Reform,” *Education Sector Reports* (May 2008), <http://all4ed.org/wp-content/uploads/2008/08/WaitingToBeWonOver.pdf>.
- ³ Jean Johnson, Ana Maria Arumi, and Amber Ott, “The Insiders: How Principals and Superintendents See Public Education Today,” *Public Agenda Reality Check* 2006, No. 4 (2006), <http://files.eric.ed.gov/fulltext/ED494314.pdf>.
- ⁴ Tom Mortenson. *Bachelor’s Degree Attainment by Age 24 by Family Income Quartiles, 1970-2009*. Oskaloosa, IA: Postsecondary Education Opportunity, 2010.
- ⁵ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2013 Mathematics and Reading Assessments, http://nationsreportcard.gov/reading_math_2013/#/student-groups.
- ⁶ Robert Balfanz et al., “Building a Grad Nation”, February 2013, <http://www.americaspromise.org/sites/default/files/legacy/bodyfiles/BuildingAGradNation2013Full.pdf>.
- ⁷ Diane Ravitch, “A Brief History of Teacher Professionalism,” prepared for the White House Conference on Preparing Tomorrow’s Teachers, March 5, 2002, <http://www2.ed.gov/admins/tchrqual/learn/preparingteachersconference/ravitch.html>.
- ⁸ “How Do You ‘Professionalize’ Teaching Within the Context of a Collective Bargaining Agreement?” *Ed In The Apple* (blog), March 20, 2014, <http://mets2006.wordpress.com/2014/03/20/how-do-you-professionalize-teaching-within-the-context-of-a-collective-bargaining-agreement/>.
- ⁹ James Coleman, U.S. Department of Education, “Equality of Educational Opportunity (COLEMAN) Study (EEOS),” 1966, <http://files.eric.ed.gov/fulltext/ED012275.pdf>.
- ¹⁰ U.S. Department of Education, National Commission on Excellence in Education, “A Nation at Risk: The Imperative for Educational Reform,” Washington: GPO, 1983, http://datacenter.spps.org/uploads/SOTW_A_Nation_at_Risk_1983.pdf.
- ¹¹ Sandra L. Stotsky, “Revising Teacher Licensing Regulations to Advance Education Reform in Massachusetts,” *Education Success Stories* (Amherst: National Evaluation Systems Inc., 2001), http://images.pearsonassessments.com/images/NES_Publications/2001_14Stotsky_466_1.pdf.
- ¹² Casey Banas, “Teacher Shortage Predicted,” *Chicago Tribune*, August 14, 1985, http://articles.chicagotribune.com/1985-08-14/news/8502220920_1_new-teachers-teacher-market-shortage.
- ¹³ Stephanie Banchemo and LeAnn Spencer, “Teacher shortage looms, study warns,” *Chicago Tribune*, December 14, 2000, http://articles.chicagotribune.com/2000-12-14/news/0012140381_1_teacher-shortage-alternative-certification-programs-assistant-superintendent.
- ¹⁴ Stephanie Banchemo and LeAnn Spencer, “Illinois teacher shortage worsens,” *Chicago Tribune*, January 17, 2002, http://articles.chicagotribune.com/2002-01-17/news/0201170161_1_teacher-shortage-illinois-classrooms-education-task-force.
- ¹⁵ See, for example: “New Mexico faces teacher shortage,” *Amarillo Globe-News*, September 23, 2000, http://amarillo.com/stories/2000/09/23/usn_newmexico.shtml; Michele Cohen, “Teacher scarcity increases shortage, could overshadow reforms,” *Sun-Sentinel*, November 10, 1985, http://articles.sun-sentinel.com/1985-11-10/news/8502190538_1_science-teacher-new-teachers-prospective-teachers; “On heels of a teacher glut, a teacher shortage,” *Christian Science Monitor*, August 30, 1983, <http://www.csmonitor.com/1983/0830/083046.html?page/2>.
- ¹⁶ American Association of State Colleges and Universities, “The Facts and Fictions about Teacher Shortages,” *Policy Matters* 5 No. 6, May 2005, <http://www.aascu.org/uploadedfiles/aascu/content/root/policyandadvocacy/policypublications/teacher%20shortage.pdf>.
- ¹⁷ U.S. Department of Education, Office of Postsecondary Education, “The Secretary’s Sixth Annual Report on Teacher Quality,” 2009, <http://www2.ed.gov/about/reports/annual/teachprep/t2r6.pdf>.
- ¹⁸ Richard Ingersoll and David Perda, “Is the Supply of Mathematics and Science Teachers Sufficient?” September 1, 2010, http://repository.upenn.edu/cgi/viewcontent.cgi?article=1229&context=gse_pubs.
- ¹⁹ Richard Ingersoll, “Do We Produce Enough Mathematics and Science Teachers?” *Kappan Magazine* 92 No. 6, March 2011,

http://www.gse.upenn.edu/pdf/rmi/Enough_Math_Teachers.pdf.

²⁰ U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), “Public School Teacher Data File,” 2003–04, https://nces.ed.gov/surveys/sass/tables/sass0304_2008338_t1n_02.asp#r1.

²¹ Stephen Sawchuk, “Colleges Overproducing Elementary Teachers, Data Find,” *Teaching the Teacher* (blog), *Education Week*, January 22, 2013, http://www.edweek.org/ew/articles/2013/01/23/18supply_ep.h32.html.

²² Bonnie Miller Rubin, “Lack of jobs leave many teachers wondering about future,” *Chicago Tribune*, August 9, 2010, http://articles.chicagotribune.com/2010-08-09/news/ct-met-schools-teaching-profession-20100809_1_male-teachers-lake-county-federation-school-year.

²³ Diane Rado, “Many are called, but few are chosen to teach,” *Chicago Tribune*, February 4, 2009, http://articles.chicagotribune.com/2009-02-04/news/0902020279_1_teacher-jobs-assistant-superintendent-elementary-school-teachers.

²⁴ Richard Ingersoll and Lisa Merrill, “Seven Trends: The Transformation of the Teaching Force,” *CPRE Working Paper*, Fall 2012, http://www.cpre.org/sites/default/files/occasionalpaper/1371_7trendsgse.pdf.

²⁵ U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), “Public Teacher Questionnaire,” 1987-88, “Public Teacher Questionnaire,” 1999-91, “Public Teacher Questionnaire,” 1993-94, “Public Teacher Questionnaire,” 1999-2000, “Public School Teacher Data File,” 2003-04, and “Public School Teacher Data File,” 2007-08, http://nces.ed.gov/surveys/sass/tables/sass0708_034_t1n.asp.

²⁶ U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), “Public School Teacher Data File,” 2003–04, https://nces.ed.gov/surveys/sass/tables/sass0304_2008338_t1n_04.asp.

²⁷ Sarah Almy and Christina Theokas, “Not Prepared for Class: High-Poverty Schools Continue to Have Fewer In-Field Teachers,” November 2010, <http://www.edtrust.org/sites/edtrust.org/files/publications/files/Not%20Prepared%20for%20Class.pdf>.

²⁸ Jennifer Woodward, “Bilingual Certification in New York State: Increasing the Number of Certified Teachers,” *NYLARNet*, Fall 2011, http://www.nylarnet.org/reports/edu_Teacher%20Certification%20Woodward.pdf.

²⁹ Florida Department of Education, “Identification of Critical Teacher Shortage Areas,” 2012, <http://www.fldoe.org/arm/pdf/ctsa1213.pdf>.

³⁰ Cynthia Leonor Garza, “Districts Seek Bilingual Ed Teachers in Mexico,” *Houston Chronicle*, February 21, 2007, <http://www.chron.com/news/houston-texas/article/Districts-look-bilingual-ed-teachers-in-Mexico-1828254.php>.

³¹ James McLeskey, Naomi Tyler, and Susan Flippin, “The Supply and Demand for Special Education Teachers: A Review of Research Regarding the Nature of the Chronic Shortage of Special Education,” October 2003, <http://copsse.education.ufl.edu/docs/RS-1/1/RS-1.pdf>.

³² U.S. Department of Education, Office of Postsecondary Education, “The Secretary’s Ninth Annual Report on Teacher Quality,” 2009, <https://title2.ed.gov/Public/TitleIIReport13.pdf>.

³³ Authors’ analysis of data from historical state-level IDEA data files, 2006-2011, <http://tadnet.public.tadnet.org/pages/712>.

³⁴ Markay L. Winston, “Guidelines for Special Education Class Size,” Amended Bulletin No. 33, Chicago Public Schools, February 15, 2013, http://www.cpsdiverselearner.org/index.php?option=com_docman&task=doc_download&cgid=658&Itemid=810.

³⁵ Douglas Harris and Tim Sass, “The Effects of NBPTS-Certified Teachers on Student Achievement,” *Calder Working Paper* No. 4, March 2007, <http://files.eric.ed.gov/fulltext/ED509659.pdf>.

³⁶ “Summary of the Improving America’s Schools Act,” *Education Week*, November 9, 1994, <http://www.edweek.org/ew/articles/1994/11/09/10asacht.h14.html?>

³⁷ “No Child Left Behind,” *Education Week*, September 19, 2011, <http://www.edweek.org/ew/issues/no-child-left-behind/>.

³⁸ Wendy Kopp (founder of Teach For America), interview by authors, via phone, December 13, 2013.

³⁹ Joel Klein (former chancellor of New York City Department of Education), interview by authors, via phone, December 19, 2013.

⁴⁰ Eric Hanushek, John F. Kain, and Steven G. Rivkin, “Teachers, Schools, and Academic Achievement,” *NBER Working Papers* 6691 (August 1998), http://www.nber.org/papers/w6691.pdf?new_window=1.

- ⁴¹ Dan Goldhaber and Dominic Brewer, “Teacher Licensing and Student Achievement,” in *Better Teachers, Better Schools*, edited by Marci Kanstoroom, and Chester E. Finn Jr, Thomas B. Fordham Foundation, 1999, http://www.edexcellencemedia.net/publications/1999/199907_betterteachersbetterschools/btrtchrs.pdf#page=91.
- ⁴² No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002), <http://www2.ed.gov/policy/elsecl/leg/esea02/pg107.html#sec9101>.
- ⁴³ Brad Jupp (policy advisor, U.S. Department of Education), interview by authors, via phone, November 27, 2013.
- ⁴⁴ Robert Gordon, Thomas Kane, and Douglas Staiger, *Identifying Effective Teachers Using Performance on the Job*, Brookings Institution, 2006, http://www.brookings.edu/views/papers/200604hamilton_1.pdf.
- ⁴⁵ Eric Hanushek, “Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data,” *The American Economic Review* 61, No. 2 (May 1971), http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%201971%20AER%2061%282%29_0.pdf.
- ⁴⁶ William L. Sanders and June C. Rivers, “Research Progress Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement,” November 1996, http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf.
- ⁴⁷ Van Schoales (chief executive officer, A+ Denver), interview by authors, via phone, December 5, 2013.
- ⁴⁸ See, for example: Raj Chetty, John N. Friedman, and Jonah E. Rockoff, “The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood,” *NBER Working Paper* 17699 (2012), http://www.nber.org/papers/w17699.pdf?new_window=1; Douglas O. Staiger and Jonah E. Rockoff, “Searching for Effective Teachers with Imperfect Information,” *Journal of Economic Perspectives* 24 No. 3 (2010), <http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.24.3.97>.
- ⁴⁹ Randi Weingarten, “Teaching and Learning over Testing,” *Huffington Post*, January 10, 2014, http://www.huffingtonpost.com/randi-weingarten/teaching-and-learning-ove_b_4575705.html.
- ⁵⁰ Kathryn M. Doherty and Sandi Jacobs, “Connect the Dots: Using Evaluations of Teacher Effectiveness to Improve Policy and Practice,” *NCTQ State of the States*, October 2013, http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report.
- ⁵¹ Jennifer L. Steele, Laura S. Hamilton, Brian M. Stecher, “Incorporating Student Performance Measures into Teacher Evaluation Systems,” *RAND Corporation technical report series*, 2010, http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR917.pdf.
- ⁵² Bill & Melinda Gates Foundation, “Ensuring Fair and Reliable Measures of Effective Teaching,” *MET Project Policy and Practice Brief*, January 2013, http://metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf.
- ⁵³ New Haven Public Schools, “Teacher Evaluation and Development—Introduction to the Process,” August 2010, http://www.nhps.net/sites/default/files/1__NHPS_TEVALDEV_Introduction_-_Aug_2010.pdf.
- ⁵⁴ Melissa Bailey, “New Eval System Pushes Out 34 Teachers,” *New Haven Independent*, September 13, 2011, http://www.newhavenindependent.org/index.php/archives/entry/new_eval_system_pushes_34_teachers_out/.
- ⁵⁵ Melissa Bailey, “In 2nd Year, Evals Ease Out 28 Teachers,” *New Haven Independent*, October 10, 2012, http://www.newhavenindependent.org/index.php/archives/entry/t-vals_year_two/.
- ⁵⁶ Melissa Bailey, “Evals Push Out 20 More Teachers,” *New Haven Independent*, February 26, 2014, http://www.newhavenindependent.org/index.php/archives/entry/new_haven_teacher_evals_3rd_year/.
- ⁵⁷ Source: Interview by authors.
- ⁵⁸ Linda Noonan (executive director, Massachusetts Business Alliance for Education), interview with authors, December 4, 2013.
- ⁵⁹ Jessica Levin and Meredith Quinn, The New Teacher Project, *Missed Opportunities*, 2003, <http://tntp.org/assets/documents/MissedOpportunities.pdf>; Jessica Levin, Jennifer Mulhern, and Joan Schunck, The New Teacher Project, *Unintended Consequences*, 2005, <http://files.eric.ed.gov/fulltext/ED515654.pdf>.
- ⁶⁰ Ibid.
- ⁶¹ Jim Blew (adviser, Walton Family Foundation), interview by authors, December 13, 2013.
- ⁶² Joel Klein (former chancellor of New York City Department of Education), interview by authors, via phone, December 19, 2013.

- ⁶³ David Weisberg et al., The New Teacher Project, *The Widget Effect*, <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.
- ⁶⁴ Ibid.
- ⁶⁵ Ibid.
- ⁶⁶ Ibid.
- ⁶⁷ District of Columbia Public Schools, “An Overview of IMPACT,” [http://www.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/An+Overview+of+IMPACT](http://www.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/An+Overview+of+IMPACT).
- ⁶⁸ Stephen Sawchuk, “D.C. Unveils Complex Evaluation System,” *Teacher Beat* (blog), *EdWeek*, October 5, 2009, http://blogs.edweek.org/edweek/teacherbeat/2009/10/dc_unveils_complex_evaluation.html.
- ⁶⁹ Sara Mead, Andrew J. Rotherham, and Rachael Brown, “The Hangover: Thinking about the Unintended Consequences of the Nation’s Teacher Evaluation Binge,” *American Enterprise Institute’s Teacher Quality 2.0 series*, September 2012, <http://bellwethereducation.org/sites/default/files/legacy/2012/09/Teacher-Quality-Mead-Rotherham-Brown.pdf>.
- ⁷⁰ Jason Kamras (chief human capital officer, District of Columbia Public Schools), interview by authors, via phone, January 15, 2014.
- ⁷¹ District of Columbia Public Schools, “IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel,” 2013, <http://dcps.dc.gov/DCPS/Files/downloads/In-the-Classroom/IMPACT%20Guidebooks/IMPACT-2013-Grp1-LR.pdf>.
- ⁷² U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2013 Mathematics and Reading Assessments, http://nationsreportcard.gov/reading_math_2013/#/state-gains.
- ⁷³ Thomas Dee and James Wyckoff, “Incentives, Selection, and Teacher Performance: Evidence from IMPACT,” *NBER Working Paper* 19529, October 2013, <http://cepa.stanford.edu/sites/default/files/w19529.pdf>.
- ⁷⁴ District of Columbia Public Schools, “IMPACTplus Guidebook.” http://dcps.dc.gov/DCPS/Files/downloads/In-the-Classroom/IMPACT%20Guidebooks/IMPACTplus_Teachers.pdf.
- ⁷⁵ Washington Teachers’ Union and District of Columbia Public Schools, “Collective Bargaining Agreement, 2007-2012,” <http://media.washingtonpost.com/wp-srv/metro/documents/teachercontract060210.pdf/>.
- ⁷⁶ Bill Turque, “D.C. teachers’ union ratifies contract, basing pay on results, not seniority,” *The Washington Post*, June 3, 2010, <http://www.washingtonpost.com/wp-dyn/content/article/2010/06/02/AR2010060202762.html>.
- ⁷⁷ Learning Point Associates, “Evaluating Teacher Effectiveness: Emerging Trends Reflected in State Phase 1 Race to the Top Applications,” May 2010, <https://www.wested.org/wp-content/uploads/lpa-evaluating-effectiveness.pdf>.
- ⁷⁸ Doherty and Jacobs, “Connect the Dots,” 2013.
- ⁷⁹ “Race to the Top Annual Performance Reports,” 2013, <https://www.rtt-apr.us>.
- ⁸⁰ Laura Sartain et al., “Rethinking Teacher Evaluation in Chicago,” November 2011, <http://ccsr.uchicago.edu/sites/default/files/publications/Teacher%20Eval%20Report%20FINAL.pdf>.
- ⁸¹ Ibid.
- ⁸² Sara Mead, “Recent State Action on Teacher Effectiveness,” August 2012, <http://bellwethereducation.org/wp-content/uploads/2012/08/RSA-Teacher-Effectiveness.pdf>.
- ⁸³ Chicago Public Schools, “REACH Students,” <http://www.cps.edu/Pages/reachstudents.aspx>.
- ⁸⁴ Susan E. Spote et al., University of Chicago Consortium on Chicago School Research, “Teacher Evaluation in Practice: Implementing Chicago’s REACH Students,” September 2013, http://ccsr.uchicago.edu/sites/default/files/publications/REACH%20Report_0.pdf.
- ⁸⁵ Ibid.
- ⁸⁶ U.S. Department of Education, “Application: Elementary and Secondary Education Act Waiver,” <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>.
- ⁸⁷ Mead, “Recent State Action,” 2012.

- ⁸⁸ Lisa Gartner and Cara Fitzpatrick, “Newly Released Numbers Underline Flaws in Florida’s Teacher Rating System,” *Tampa Bay Times*, December 3, 2013, <http://www.tampabay.com/news/education/k12/newly-released-numbers-underline-flaws-in-floridas-teacher-rating-system/2155404>.
- ⁸⁹ U.S. Department of Education, “Hillsborough County Public Schools,” prepared for the Labor-Management Collaboration, <http://www.ed.gov/labor-management-collaboration/conference/hillsborough-county-public-schools>.
- ⁹⁰ Florida Department of Education, “Personnel Evaluation Data for Classroom Teachers by District,” January 18, 2013, <http://www.fldoe.org/profdev/pdf/StatewideResults.pdf>.
- ⁹¹ Venessa A. Keesler and Carla Howe, Michigan Department of Education, “Understanding Educator Evaluations in Michigan, Results from Year 1 of Implementation,” November 2012, http://www.michigan.gov/documents/mde/Educator_Effectiveness_Ratings_Policy_Brief_403184_7.pdf.
- ⁹² Tennessee Department of Education, “Teacher Evaluation in Tennessee: A Report on Year 1 of Implementation,” July 2012, http://www.tn.gov/education/doc/yr_1_tchr_eval_rpt.pdf.
- ⁹³ Stephen Sawchuk, “Teacher Evaluation Sparks Clash in Pittsburgh,” *Education Week*, January 22, 2014, <http://www.edweek.org/ew/articles/2014/01/22/18pittsburgh.h33.html>.
- ⁹⁴ Data from the Office of Human Capital, District of Columbia Public Schools. Cited with permission.
- ⁹⁵ Lisa Watts, Joy Singleton-Stevens, and Mark Teoh, “Lessons from the Leading Edge: Teachers’ View on the Impact of Evaluation Reform,” 2013, http://www.teachplus.org/uploads/Documents/1371738773_lessons_from_the_leading_edge.pdf.
- ⁹⁶ Rebecca Harris, “New teacher evaluations get positive reviews,” *Catalyst Chicago*, September 18, 2013, <http://www.catalyst-chicago.org/notebook/2013/09/18/60249/new-teacher-evaluations-get-positive-reviews>.
- ⁹⁷ Spote et al., “REACH Students,” 2013
- ⁹⁸ Susan E. Spote et al., University of Chicago Consortium on Chicago School Research, “Teacher Evaluation in Practice: Implementing Chicago’s REACH Students,” September 2013, http://ccsr.uchicago.edu/sites/default/files/publications/REACH%20Report_0.pdf.
- ⁹⁹ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2013 Mathematics and Reading Assessments, http://nationsreportcard.gov/reading_math_2013/#/state-gains.
- ¹⁰⁰ Donald Boyd et al., “The Effect of Certification and Preparation on Teacher Quality,” *Future of Children* 17 No. 1 (Spring 2007), https://www.princeton.edu/futureofchildren/publications/docs/17_01_03.pdf.
- ¹⁰¹ Julie Greenberg, Arthur McKee, and Kate Walsh, NCTQ, “Teacher Prep Review,” December 2013, http://www.nctq.org/dmsView/Teacher_Prep_Review_2013_Report.
- ¹⁰² Tim Daly (president, TNTP), interview with authors, via phone, December 6, 2013.
- ¹⁰³ Chad Aldeman et al., “A Measured Approach to Improving Teacher Preparation,” *Education Sector Policy Briefs* 2011, http://www.educationsector.org/sites/default/files/publications/TeacherPrep_Brief_RELEASE.pdf.
- ¹⁰⁴ Mead, Rotherham, and Brown, “The Hangover,” 2012; Gordon, Kane, and Staiger, “Identifying Effective Teachers,” 2006; Donald Boyd et al., “How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement,” *NBER Working Paper* 11844, December 2005, <http://www.nber.org/papers/w11844>.
- ¹⁰⁵ Greenberg, McKee, and Walsh, “Teacher Prep Review”, 2013.
- ¹⁰⁶ Jackie Mader, “Alternative routes to teaching become more popular despite lack of evidence,” *The Hechinger Report*, May 16, 2013, http://hechingerreport.org/content/alternative-routes-to-teaching-become-more-popular-despite-lack-of-evidence_12059/.
- ¹⁰⁷ Kate Walsh and Sandi Jacobs, NCTQ, “Alternative Certification Isn’t Alternative,” September 2007, <http://files.eric.ed.gov/fulltext/ED498382.pdf>.
- ¹⁰⁸ See, for example: P.T. Decker, D.P. Mayer, and S. Glazerman. “The Effects of Teach For America on Students: Findings from a National Evaluation.” Princeton, NJ: Mathematica Policy Research, Inc., June 2004, <http://www.mathematica-mpr.com/publications/pdfs/teach.pdf>.; Melissa A. Clarke et al., “The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs,” September 2013,

- <http://ies.ed.gov/ncee/pubs/20134015/pdf/20134015.pdf>; Strategic Data Project, Harvard University, “Los Angeles Unified School District,” *SDP Human Capital Diagnostic*, November 2012, <http://www.gse.harvard.edu/cepr-resources/files/news-events/sdp-laUSD-hc.pdf>; George H. Noell and Kristin Gansle, “Teach For America’s Teachers’ Contributions to Student Achievement in Louisiana in Grades 4-9: 2004-2005 to 2006-2007,” October 27, 2009, http://www.nctq.org/docs/TFA_Louisiana_study.PDF; Tennessee Higher Education Commission, State Board of Education, “2013 Report Card on the Effectiveness of Teacher Training Programs,” November 1, 2013, http://www.tn.gov/thec/Divisions/fttt/13report_card/1_Report%20Card%20on%20the%20Effectiveness%20of%20Teacher%20Training%20Programs.pdf; Margaret Raymond, Stephen Fletcher, and Javier Luque, “Teach For America: An Evaluation of Teacher Differences and Student Outcomes in Houston, Texas,” *CREDO Report*, July 2001, <http://credo.stanford.edu/downloads/tfa.pdf>; Herbert M. Turner et al., “Evaluation of Teach For America in Texas Schools,” December 2012, http://www.edvanceresearch.com/images/Evaluation_of_Teach_For_America_in_Texas_Schools_1-2013.pdf.
- ¹⁰⁹ See, for example: Jason Richwine and Andrew G. Biggs, “Assessing the Compensation of Public School Teachers,” November 1, 2011, http://www.aei.org/files/2011/11/02/-assessing-the-compensation-of-publicschool-teachers_19282337242.pdf; and Secretary Duncan’s response: Arne Duncan, “Teacher Pay Study Asks the Wrong Question, Ignores Facts, Insults Teachers,” *The Huffington Post*, November 9, 2011, http://www.huffingtonpost.com/arne-duncan/teacher-pay-study-asks-th_b_1084881.html.
- ¹¹⁰ Gordon, Kane, and Staiger, “Identifying Effective Teachers,” 2006.
- ¹¹¹ District of Columbia Public Schools, “IMPACTplus Guidebook,” http://dcps.dc.gov/DCPS/Files/downloads/In-the-Classroom/IMPACT%20Guidebooks/IMPACTplus_Teachers.pdf.
- ¹¹² Dee and Wyckoff, “IMPACT,” 2013.
- ¹¹³ Elena Silva, “ProComp Evaluation Results,” *The Quick & the Ed* (blog), October 31, 2011, <http://www.quickanded.com/2011/10/procomp-final-evaluation-results.html?>
- ¹¹⁴ Gary W. Ritter et al., “Year Two Evaluation of the Achievement Challenge Pilot Project in the Little Rock Public School District,” January 22, 2008, http://uark.edu/ua/der/Research/merit_pay/year_two/Full_Report_without_Appendices.pdf.
- ¹¹⁵ Matthew G. Springer et al., “Final Report, Experimental Evidence from the Project on Incentives in Teaching (POINT),” September 30, 2012, <https://my.vanderbilt.edu/performanceincentives/files/2012/09/Final-Final-Report-Experimental-Evidence-from-the-Project-on-Incentives-in-Teaching-2012.pdf>.
- ¹¹⁶ Talia Milgrom-Elcott (former program officer, Carnegie Corporation of New York), interview by authors, via phone, December 17, 2013.
- ¹¹⁷ Arne Duncan (U.S. Secretary of Education), interview by authors, via phone, December 11, 2013.
- ¹¹⁸ Matthew Schatz, “In New York City, Building Upon the Past,” *Eduwonk* (blog), July 25, 2013, <http://www.eduwonk.com/2013/07/in-new-york-city-building-upon-the-past.html>.
- ¹¹⁹ Washington Teachers’ Union and District of Columbia Public Schools, “Collective Bargaining Agreement, 2007-2012,” <http://media.washingtonpost.com/wp-srv/metro/documents/teachercontract060210.pdf/>.
- ¹²⁰ Joel Klein (former chancellor of New York City Department of Education), interview by authors, via phone, December 19, 2013.
- ¹²¹ Jal Mehta and Steven Teles. “Professionalization 2.0: The Case for Plural Professionalization in Education,” *American Enterprise Institute’s Teacher Quality 2.0 series*, September 12, 2013. Cited with permission.
- ¹²² Scholastic and Bill & Melinda Gates Foundation, “Primary Sources: America’s Teachers on Teaching in an Era of Change,” 3rd edition, 2013, <http://www.scholastic.com/primarysources/PrimarySources3rdEditionWithAppendix.pdf>.
- ¹²³ U.S. Department of Education, “Findings from the 2012-2013 Survey on the Use of Funds under Title II,” June 2013, <http://www2.ed.gov/programs/teacherqual/resources.html>.
- ¹²⁴ Government Accountability Office, “K-12 Education: Characteristics of the Investing in Innovation Fund,” February 7, 2014, <http://www.gao.gov/assets/670/660743.pdf>.
- ¹²⁵ Allison Gulamhussein, Center for Public Education, “Teaching the Teachers,” September 2013, <http://www.centerforpubliceducation.org/Main-Menu/Staffingstudents/Teaching-the-Teachers-Effective-Professional-Development-in-an-Era-of-High-Stakes-Accountability/Teaching-the-Teachers-Full-Report.pdf>.
- ¹²⁶ Kwang Suk Yoon et al., “Reviewing the Evidence on how Teacher Professional Development Affects Student Achievement,” *REL Southwest* No. 33, October 2007, http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033.pdf.

- ¹²⁷ Linda Darling-Hammond et al., “Professional Learning in the Learning Profession,” February 2009, <http://learningforward.org/docs/pdf/nsdcstudy2009.pdf>.
- ¹²⁸ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, “Middle School Mathematics Professional Development Impact Study Findings after the Second Year of Implementation,” May 2011, <http://ies.ed.gov/ncee/pubs/20114024/index.asp>.
- ¹²⁹ U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, “Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement,” September 2008, http://ies.ed.gov/ncee/pubs/20084030/summ_e.asp.
- ¹³⁰ Rebecca Perry and Catherine Lewis, “Improving the Mathematical Content Base of Lesson Study: Summary of Results,” September 1, 2011, <http://www.lessonresearch.net/IESAbstract10.pdf>.
- ¹³¹ Sporte et al., “REACH Students,” 2013.
- ¹³² Scholastic and Bill & Melinda Gates Foundation, “Primary Sources,” 2013.
- ¹³³ Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, “How and Why Do Teacher Credentials Matter for Student Achievement,” *CALDER* 2, March 2007, <http://files.eric.ed.gov/fulltext/ED509655.pdf>.
- ¹³⁴ See, for example: Steven G. Rivkin, Eric A. Hanushek, and John F. Kain, “Teachers, Schools, and Academic Achievement,” *Econometrica* 73, No. 2, March 2005, <http://www.econ.ucsb.edu/~jon/Econ230C/HanushekRivkin.pdf>; Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, “Teacher-Student Matching and the Assessment of Teacher Effectiveness,” *NBER Working Paper* 11936, January 2006, http://www.nber.org/papers/w11936.pdf?new_window=1.
- ¹³⁵ TNTP, “The Irreplaceables,” 2012, http://tntp.org/assets/documents/TNTP_Irreplaceables_2012.pdf.
- ¹³⁶ Kati Haycock (president, The Education Trust), interview by authors, via phone, December 20, 2013.
- ¹³⁷ Eric Hanushek (Paul and Jean Hanna Senior Fellow, Hoover Institution), interview by authors, via phone, December 2, 2013.
- ¹³⁸ Jack Jennings (founder, Center on Education Policy), interview by authors, via phone, December 16, 2013.
- ¹³⁹ Arne Duncan (U.S. Secretary of Education), interview by authors, via phone, December 11, 2013.
- ¹⁴⁰ Julie Greenberg, Laura Pomerance, and Kate Walsh, NCTQ, “Student Teaching in the United States,” July 2011, http://www.nctq.org/dmsView/Student_Teaching_United_States_NCTQ_Report.
- ¹⁴¹ Donald Boyd et al., “Teacher Preparation and Student Achievement,” *CALDER* 20, August 2008, http://www.urban.org/UploadedPDF/1001255_teacher_preparation.pdf.
- ¹⁴² Gordon, Kane, and Staiger, “Identifying Effective Teachers,” 2006.
- ¹⁴³ Jack Jennings (founder, Center on Education Policy), interview by authors, via phone, December 16, 2013.
- ¹⁴⁴ NCTQ, “Teacher Layoffs: Rethinking ‘Last Hired, First Fired’ Policies,” February 2010, http://www.nctq.org/dmsView/Teacher_Layoffs_Rethinking_Last-Hired_First-Fired_Policies_NCTQ_Report.
- ¹⁴⁵ Mead, “Recent State Action,” 2012.
- ¹⁴⁶ TNTP, “The Irreplaceables,” 2012.
- ¹⁴⁷ Disclosure: Bellwether consulted for New Classrooms and Rocketship Education.
- ¹⁴⁸ See, for example: RAND Corporation, “Does an Algebra Course Improve Student Learning?” 2013, http://www.rand.org/content/dam/rand/pubs/research_briefs/RB9700/RB9746/RAND_RB9746.pdf.
- ¹⁴⁹ Center for Responsive Politics, “Heavy Hitters: Top All-Time Donors, 1989-2014,” <https://www.opensecrets.org/orgs/list.php>.
- ¹⁵⁰ Melissa Bailey, “A teachers union embraces reform in New Haven, creating a model for others,” *The Hechinger Report*, June 11, 2013, http://hechingerreport.org/content/a-teachers-union-embraces-reform-in-new-haven-creating-a-model-for-others_11744/.