

## Federal Efforts to Improve Teacher Quality

Matthew A. Kraft

In 1981, newly appointed Secretary of Education Terrel H. Bell convinced President Ronald Reagan to appoint the National Commission on Excellence in Education. The commission's landmark report, *A Nation at Risk*, first planted the seeds for a direct federal role in shaping the teaching profession in the United States. The report identified the poor skills and training of the current teacher workforce as one of four core causes of the "rising tide of mediocrity" in US public education. At the same time, the report warned that the "professional working life of teachers is on the whole unacceptable," and that serious shortages exist in key STEM fields. This call to arms was met with skepticism by President Reagan, who was committed to shrinking the already limited role of the federal government in education. Two decades later, however, Presidents George W. Bush and Barack Obama would answer this call by advancing major federal policy initiatives aimed at improving the quality of public school teachers in the United States.

Bush's and Obama's federal education reforms were remarkably similar in their goals and ambitions. Bush's No Child Left Behind (NCLB) Act and Obama's Race to the Top (RTTT) and NCLB state waiver programs leveraged federal funding and authority to address four broad areas: academic standards, data and accountability, teacher quality, and school turnarounds. This chapter focuses specifically on how these efforts have influenced the teaching profession. During Bush's and Obama's combined sixteen years in office, the federal government succeeded in fundamentally changing licensure requirements and evaluation systems for public school teachers. Reflecting on the successes and failures of these reforms provides important lessons about the potential and limitations of federal policy as a tool for improving the quality of the US teacher workforce.

## TEACHER LICENSURE REFORMS UNDER BUSH

*A Nation at Risk* elevated concerns that many teachers were not well prepared or qualified to teach their subject matter. The report asserted, “Half of the newly employed mathematics, science, and English teachers are not qualified to teach these subjects.” The creation of the Schools and Staffing Survey (SASS)—a federally funded national survey of teachers, first administered in 1988—helped to bring this picture of poorly qualified and out-of-field teaching further into focus. Richard Ingersoll’s analyses of the SASS data documented that a third of math teachers, a quarter of English teachers, and a fifth of science and social studies teachers did not have a degree in their subject.<sup>1</sup>

Research at the time also demonstrated that states held widely different and often low standards for obtaining a licensure to teach in public schools.<sup>2</sup> Unlike many countries, where authority over public education is centralized at the national level, states in the United States have sole authority to establish licensure criteria and grant licenses to teach in public schools. When Bush took office, less than half of all states required high school teachers to have majored in their subject area.<sup>3</sup> Even among the twenty-nine states that required teaching candidates to take a subject matter test, minimum scores were often set so low that nearly everyone passed.<sup>4</sup> States also commonly offered emergency or temporary licensures. In 2001, 6 percent of teachers nationally lacked full certification and in nine states more than 10 percent of teachers were uncertified.<sup>5</sup>

### NCLB’s Highly Qualified Teacher Provisions

President Bush signed NCLB, a reauthorization of the Elementary and Secondary Education Act (ESEA), into law on January 8, 2002. The 670-page document included a set of Highly Qualified Teacher (HQT) provisions that were the result of a bipartisan compromise between the Republican administration and congressional Democrats led by Senator Ted Kennedy. The HQT provisions were intended to establish a set of standard licensure requirements across states in order for teachers to be considered high quality.<sup>6</sup> The law defined “highly qualified” as meeting three broad requirements: 1) possessing a bachelor’s degree from a four-year institution, 2) holding a full state certification, and 3) demonstrating subject area competence.

New entrants into the teaching profession were held to the highest standards. The law stipulated that to be considered an HQT, new teachers had

to demonstrate subject matter knowledge and competence in teaching skills. New elementary school teachers were required to pass a state examination covering reading, writing, math, and other areas of the basic elementary curriculum. New middle and high school teachers were given four options to demonstrate competence in their subject area: pass a state exam in their content area, earn an undergraduate or graduate degree in their subject area, accumulate the equivalent coursework of an undergraduate degree in their subject area, or attain an advanced certificate or credential. The law gave states greater discretion for certifying current teachers as highly qualified. Like new teachers, current teachers could meet HQT requirements by passing subject matter tests or obtaining the equivalent of an undergraduate degree in their subject. States were also allowed to certify teachers against a locally created “high, objective, uniform state standard of evaluation” (HOUSSE).

Lawmakers established an aggressive timeline for states to comply with these standards and provided federal funding to support implementation. All new teachers and paraprofessionals hired with Title I funds were required to meet HQT standards for the start of the 2002–2003 school year. These requirements were expanded to all core-subject teachers and paraprofessionals by 2005–2006. Total federal spending by the US Department of Education (ED) increased by 46 percent between 2002–2003 and 2005–2006 when the HQT mandate came into effect.<sup>7</sup> In fiscal year 2002 alone, more than \$3 billion was made available as formula grants to states and local education agencies (LEAs) to be used for professional development, developing alternative routes to certification, helping teachers meet certification requirements, recruiting and retaining HQTs, testing teachers in subject areas, and creating programs for HQTs.<sup>8</sup>

The law also made several enforcement mechanisms available to the Bush administration. Six months after the law was passed, states were required to submit reports to ED with details of how they would meet HQT requirements. States then had to submit an annual report card detailing the professional qualifications of public school teachers as well as the percentage of teachers with emergency or provisional credentials. The law also included language requiring state plans to “ensure that poor and minority children are not taught at higher rates than other children by inexperienced, unqualified, or out-of-field teachers” and to disaggregate these statistics across schools in the top and bottom quintiles of poverty in the state.<sup>9</sup> As a core compliance mechanism, the federal government could choose to withhold Title

I funds from states that failed to meet these requirements. It also required schools to notify parents when their children had been taught for four or more consecutive weeks by teachers who were not highly qualified.

### *The effects of the HQT provisions*

The core motivating principle behind the HQT provisions was both intuitive and widely appealing: teachers should know the material they teach. A national poll conducted by *Phi Delta Kappan* in 2002 found that 96 percent of respondents thought that teachers in the public schools in their community should be required to take a statewide competency test in the subjects they teach.<sup>10</sup> Examples of similar policies requiring potential primary and secondary teachers to pass an examination to enter the teaching profession can be found in fifteen of the forty-one member countries in the Organisation for Economic Co-operation and Development (OECD).

The HQT provisions resulted in some clear, if modest, successes. The percent of teachers teaching on state waivers dropped from 3.1 in 2003–2004 to 1.4 in 2007–2008.<sup>11</sup> The law also compelled states and districts to build new data systems tracking public school teachers' qualifications, which created an important new data source for examining the distribution of teacher quality. Finally, the law induced testing companies such as Educational Testing Service (ETS) to develop a new generation of tests to assess the basic knowledge, content expertise, and pedagogical skills of teachers.

These successes were undercut in many ways by ambiguity in the law, states' focus on compliance, limited federal enforcement, and unintended consequences. From the outset, the HQT provisions reflected a lack of understanding about the certification process and differences in local teacher labor markets. For example, states differed in whether they required middle school teachers to obtain an elementary or secondary teacher licensure. The degree to which teachers had to demonstrate content knowledge in their specific courses or their subject area more broadly was also unclear. Did high school science teachers who taught multiple subjects have to pass tests in chemistry, biology, and physics? Secretary of Education Rod Paige ultimately issued new guidance on certification flexibility for middle school teachers and high school science teachers, but never provided guidance on the certification requirements for special education teachers to be considered highly qualified.

Demonstrating content mastery for current middle school and high school teachers became an exercise in compliance that alienated teachers

and did little to strengthen their skills. States overwhelmingly used their flexibility under the HOUSSE provisions to provide certified teachers avenues through which they could meet HQT standards in the course of their current professional duties. For example, certified teachers could meet the HQT standards by serving on a curriculum committee, mentoring a new teacher, leading an academic club, completing a National Board application, or attending conferences.<sup>12</sup> In many states, the bar was set so low on new state licensure and subject matter exams that nearly all teachers passed. Nationally, 96 percent of teachers passed all tests they took for their areas of specialization from 2003–2004 through 2008–2009.<sup>13</sup>

There were few, if any, actual consequences for states that circumvented the HQT standards through lax HOUSSE rules or that failed to make progress toward the goal of 100 percent of teachers being highly qualified by the end of 2005–2006.<sup>14</sup> It is not evident that Secretaries Page or Spellings ever put a state on notice or warned a state of the potential loss of federal funding under Title I for failing to comply with the HQT provisions. States skirted even the parental Right-to-Know requirements by interpreting them as applying only to parents who requested HQT information, or by classifying only substitute teachers as not being highly qualified.<sup>15</sup>

Neither the executive branch nor ED focused attention on the HQT equity provisions that required states to report on and address inequities in the distribution of highly qualified teachers. These provisions were largely the product of bipartisan compromise, which allowed Congressman George Miller (D-CA) to include the language in the law. It was not until 2006 that ED even asked states to produce equity plans. The vast majority of these reports did little more than provide data found in state annual reports, lacking any targeted strategies for addressing the inequitable distribution of teachers.<sup>16</sup>

As a whole, there is very little evidence to suggest that the HQT provisions overall, or licensure exams specifically, have increased the quality of the teacher workforce. Licensure test scores are positively correlated with teachers' contributions to student learning, but only weakly so.<sup>17</sup> Most evidence suggests that the additional barriers to becoming a public school teacher imposed by licensure exams did more to dissuade higher-quality teachers from entering the profession than to keep less qualified teaching candidates out.<sup>18</sup> Licensure tests also created new barriers to important efforts to diversify the teacher workforce given that minority teaching candidates passed the exams at lower rates, on average.<sup>19</sup>

## TEACHER EVALUATION REFORM UNDER OBAMA

The policy conversation about teacher quality evolved rapidly during President Bush's two terms in office. A growing body of empirical research on teachers emerged in this period that exploited new district administrative datasets linking students to their teachers. Three seminal findings served as signposts for the signature education reforms initiatives of President Obama and his Secretary of Education, Arne Duncan: 1) the effects teachers have on student learning are large and vary considerably across teachers; 2) observable teacher characteristics such as those on résumés are at best weakly related to teachers' effects on student learning; and 3) teacher evaluation systems failed to differentiate among teachers despite large differences in effectiveness.

Research on teacher effects by William Sanders and June Rivers at the University of Tennessee's Value-Added Research and Assessment Center first captured the attention of policy makers at a national scale in 1996. Sanders and Rivers documented that a student who was taught by three consecutive top-quintile teachers gained over 50 percentile points more on state tests than peers taught by three consecutive bottom-quintile teachers.<sup>20</sup> Their research, along with that of Eric Hanushek, Tom Kane, and their colleagues, generated compelling evidence for what has become a stylized fact in education circles—teachers are the single largest school-based determinant of student achievement.<sup>21</sup> Researchers and policy makers interpreted these findings as evidence that ensuring students were taught by the best teachers several years in a row was enough to close racial and socioeconomic achievement gaps.

At the same time, new evidence cast doubt on input-based policies such as NCLB's HQT provisions. Studies consistently found only weak relationships between teachers' educational attainment, course taking, cognitive ability, or licensure type and their students' achievement gains.<sup>22</sup> Given these measurement challenges, policy makers saw new value-added models as a way to directly measure teacher output and better inform human resource decisions in schools.<sup>23</sup> School chancellors Joel Klein and Michelle Rhee were two national figures in the vanguard of this movement to reform district human capital practices in New York City and Washington, DC, public schools, respectively. Private foundations and philanthropic organizations, such as the Gates and Broad Foundations and the New Schools Venture Fund, invested millions of dollars to support these efforts in districts across the country. The release of *The Widget Effect* by The New Teacher Project

(TNTP) in June 2009 galvanized national attention on the failure of public schools to recognize and respond to differences in teacher effectiveness. The report used administrative data from twelve districts to document that nearly every teacher received a satisfactory rating.<sup>24</sup>

The growing momentum toward greater teacher accountability in the United States did not reflect broader international trends. The use of high-stakes evaluations as an approach to improving the quality of the teacher workforce varies considerably across countries. Several high-performing countries (e.g., Finland) place little emphasis on evaluations, while others (e.g., Singapore, Japan, and Korea) invest in teacher evaluation with a focus on informing teacher promotion and compensation.<sup>25</sup> Very few countries incorporate performance measures based on student test scores, relying instead on classroom observations by supervisors and often informal peer accountability.<sup>26</sup>

### **Race to the Top**

A month after taking office, President Obama signed the American Recovery and Reinvestment Act (ARRA). While the bulk of the \$115 billion for education funding was allocated to cash-strapped states and districts to help them avoid layoffs, \$4.35 billion was allotted for a competitive state grants program. This program, known as Race to the Top, became the central education initiative of Obama's first term. The law required ED to award the full \$4.35 billion in less than two years given that ARRA's purpose was to stimulate the economy.

The grant guidelines released by ED in November 2009 provided states with a detailed five-hundred-point rubric on which their applications would be judged. Points were allotted across four key areas: 1) Standards and Assessments, 2) Data Systems to Support Instruction, 3) Great Teachers and Leaders, and 4) Turning Around Struggling Schools. Across these four areas, Great Teachers and Leaders was given the most weight, with 27.6 percent of all possible points. Within this category, points were allocated across five reform areas: teacher evaluation, the equitable distribution of teachers, traditional teacher preparation, alternative teacher preparation, and professional development. Teacher evaluation criteria comprised the largest share, with 58 out of the 138 points.

The RTTT application rubric detailed specific evaluation system features that the Obama administration incentivized states to adopt. States could earn five points for measuring student achievement growth, fifteen points

for implementing rigorous evaluation systems that incorporated student achievement growth and used multiple rating categories, ten points for conducting annual evaluations and providing feedback, and twenty-eight points for using evaluations to make decisions about professional development, tenure and promotion, compensation, and retention. Although these elements constituted only 11.6 percent of the total points possible, states passed new legislation overturning state laws and collectively bargained agreements that had previously prevented districts from linking teachers to their students' test scores.

Few states passed on the opportunity to secure hundreds of millions of dollars in additional federal grant aid during the Great Recession. Forty states, along with the District of Columbia, submitted applications to the first round of the RTTT competition in January 2010, just two months after the request for proposals was released. Some states contracted with outside consulting firms, such as Education First and McKinsey & Company, to draft their proposals within the short timeframe.<sup>27</sup> ED awarded grants only to Delaware and Tennessee, but encouraged applicants to reapply in June. Out of thirty-six applicants to the second round, nine states and DC were announced as winners in September 2010. The Obama administration succeeded in securing additional funds for the 2011, 2012, and 2013 fiscal years to conduct a third state grant competition and two rounds of district grant competitions. Seven additional states and twenty-one districts won competitive grants in these rounds. Across forty-six states and the District of Columbia, 615 applications representing over 1,689 districts entered the RTTT district grant competition.<sup>28</sup> In total, ED awarded \$4.63 billion to state and district competition winners.<sup>29</sup>

The RTTT competition provided the Obama administration with a powerful lever for inducing states to commit to rapid and sweeping teacher evaluation reforms. This was not the first time the federal government leveraged competitive grant programs to advance education reforms. For example, Congress established the Teacher Incentive Fund (TIF) in 2006 under Bush. TIF supported district incentive pay proposals designed to recruit effective teachers to high-need schools and reward teachers that received strong performance evaluations. Unlike during the Bush administration, the Department of Education under Obama used all available means through the RTTT grant competition to compel states to enact its proposed reforms. Moreover, Secretary Duncan and ED were strategic in their scoring of grant



applications—they required most states to revise and resubmit applications if proposals were not closely aligned with the administration’s vision of high-stakes teacher evaluation reforms. ED also actively monitored the progress of RTTT winners by requiring annual performance reports, conducting site-based visits in struggling states, and placing states on “high-risk” status if they did not meet the commitments in their grant proposals in a timely manner.<sup>30</sup> ED even temporarily withheld funding for Hawaii, Georgia, and Maryland because these states had not secured the support of state unions or LEAs in adopting new teacher evaluation systems.<sup>31</sup>

#### *State waivers from NCLB*

While the Obama administration oversaw the RTTT competition, states and schools continued to operate under the accountability structure created by NCLB. The share of schools failing to meet Adequate Yearly Progress (AYP) goals began to increase by double digits as the mandate of 100 percent proficiency in math and reading by 2014 approached. In 2011, 48 percent of schools did not meet AYP. Secretary Duncan warned that this number would surpass 80 percent unless the federal government took action.<sup>32</sup> The consensus on Capitol Hill was that NCLB was not working, but Congress’s inability to reauthorize ESEA meant that schools were increasingly facing sanctions under the law.

Under Secretary Duncan’s guidance, ED invoked section 9401 of NCLB to grant states flexibility in complying with the law. The administration offered states relief from NCLB’s requirements in exchange for a commitment to pursue four education reform principles: 1) college- and career-ready standards, 2) state accountability systems for schools, 3) teacher evaluation reforms, and 4) reduced reporting requirements for schools and districts. ED’s guidance and expectations around teacher evaluation reforms were a clear continuation of the priorities spelled out in the RTTT competition. States had to commit to adopting teacher evaluation systems that differentiated across at least three performance levels, using multiple measures to determine performance (including weighing student academic growth as a “significant factor”), evaluating teachers on a regular basis, providing clear and timely feedback, and informing personnel decisions for teachers using performance measures. Unlike NCLB and the RTTT grant competition, no federal funding was made available to support states in implementing the reforms required for receiving a waiver.

The Obama administration's approach to state waivers marked an important transition from using strong financial incentives under RTTT to leveraging the threat of increasingly burdensome sanctions under NCLB. States were not required to apply for a waiver, but the alternative required them to implement an increasingly acute set of corrective actions when schools failed to meet AYP multiple years in a row. This included setting aside Title I funding for tutoring programs and restructuring failing schools by dismissing the school administration and most of the staff. Between February 2012 and April 2014, forty-three states and DC were granted a waiver from NCLB's provisions.

As with RTTT, ED under Obama actively enforced states' commitments to the four principles outlined in their NCLB waivers. Several states, including Kansas and Oregon, were placed on high-risk status, while Washington's waiver was revoked for failing to link teacher performance evaluations to student growth measures. The waiver process continued to push states to adopt and implement high-stakes teacher evaluation systems when public support for such systems was eroding. From 2012 to 2014, public support for using student performance on standardized tests as part of teacher evaluations had dropped from 52 percent to 38 percent.<sup>33</sup>

#### *The effect of RTTT and state waivers on teacher quality*

In many ways, the Obama administration succeeded in transforming the way in which teachers are evaluated in the United States. When Obama took office, teacher evaluation was an infrequent compliance exercise, with few, if any, consequences tied to teachers' summative ratings. By the end of Obama's second term, all but six states had radically reformed their teacher evaluation systems to incorporate multiple performance measures, often including value added to student achievement; to rate teachers on a scale with multiple categories; and, in some cases, to tie these ratings to decisions related to professional development, retention, and tenure.

These reforms resulted in several positive developments. Building on progress made under NCLB, states that won RTTT awards, and many that did not, invested in next-generation administrative data systems that provide a wealth of information to inform policy and practice. States also abandoned the binary checklists that assessed teachers on professional standards related only tangentially to instruction, such as the cleanliness and organization of the classroom. Administrators now evaluate teachers with locally

adapted versions of research-based observation rubrics that focus on instructional quality and teacher-student interactions. These rubrics provide clear standards and a common language for discussing high-quality instructional practice.

At the same time, these new evaluation systems fell short of reformers' ambitious goals. Teacher evaluation reforms were championed as teacher improvement systems that would be used to identify and remove low performing teachers, support teachers' professional growth through observation and feedback, and attract and retain high performing teachers with performance pay. Although the proportion of teachers rated in categories below proficient increased to just over 4 percent, on average, the number of teachers rated in the lowest category, unsatisfactory, remained less than 1 percent under the new systems.<sup>34</sup> Interviews with principals revealed a variety of reasons for this, including the amount of time it takes to formally document unsatisfactory performance and concerns that schools will be unable to fill open positions with stronger candidates.

The new evaluation systems did succeed, however, at differentiating teacher performance at the top end of the distribution. Across states, an average of 37 percent of teachers were identified as highly effective, with considerable variation across individual states. Publicly recognizing exemplary practice is an important and nontrivial change to a professional culture that has eschewed differentiation. However, states rarely linked these new high-performance ratings to teacher compensation or opportunities for professional advancement. Only ten states adopted some type of merit pay program.<sup>35</sup> Even districts that were awarded grants to implement merit pay systems through ED's Teacher Incentive Fund commonly designed bonus systems with very low standards and small payouts, which served as de facto increases in base salary for most teachers.

Most RTTT applications and state rollouts of the new evaluation systems framed them as tools for promoting teacher development. Evidence has shown that rigorous evaluation supported by frequent observation and feedback can improve teacher practice and student achievement.<sup>36</sup> However, few states or districts funded or implemented systems capable of supporting teacher professional growth. Some states, particularly those that were compelled to adopt evaluation reforms in exchange for an NCLB waiver, lacked the buy-in and funding necessary to move beyond minimum compliance. Others tried but ran into substantial implementation challenges in scaling

new evaluation systems. Because hiring outside evaluators or coaches was often prohibitively costly, districts gave administrators the added responsibility of conducting observations and providing feedback to teachers. This meant that administrators had to provide instructional feedback to teachers across grade levels and content areas irrespective of whether they had prior experience teaching these grades and subjects. Administrators had little time for conducting more than the minimum number of required observations, and post-observation conferences became infrequent and largely a formality.<sup>37</sup>

Recent research also reveals that high-stakes evaluation reforms decreased the number of new entrants into the teaching profession and increased turnover among current classroom teachers. On one hand, this is concerning given teacher shortages in certain regions and subject areas as well as the negative effects associated with turnover on both schools and students. Studies show it costs between \$2,000 and \$8,000 to replace a teacher, and that turnover undercuts efforts to coordinate instruction and lowers student achievement.<sup>38</sup> At the same time, if evaluation reforms made teaching more attractive to high performing current and prospective teachers but less attractive to low performing teachers, then these patterns could result in a net gain in the quality of the teacher workforce. Recent research has found that evaluation reforms discouraged would-be teachers with stronger and weaker qualifications equally.<sup>39</sup> However, several studies have found that turnover among current teachers has been disproportionately concentrated among low performing teachers.<sup>40</sup>

Progress on some elements of Obama's teacher evaluation reform agenda began to erode even before his second term ended. In an attempt to relieve pressure on states, ED created a waiver amendment process to request additional time for pilot-testing the new systems. Still, the number of states committed to including student growth measures as part of the evaluation system dropped from forty-three states in 2015 to thirty-nine states in 2017.<sup>41</sup> Moreover, the number of states requiring student growth to be the preponderant measure in teachers' overall ratings decreased from nineteen in 2013 to ten in 2016.<sup>42</sup> Political opposition to the Obama administration's use of NCLB waivers to advance teacher evaluation reforms culminated in the passage of the Every Student Succeeds Act (ESSA) in December 2015, which explicitly prohibited the Department of Education from mandating evaluation requirements as a condition of Title I funding or state waivers.

## LESSONS LEARNED

### **Lesson 1: Federal Data Collection and State Reporting Requirements Are Valuable Practices**

The federal government was responsible for the development of the data that, in many ways, shaped efforts to improve education policy under Bush and Obama. For example, the federally funded Schools and Staffing Survey produced compelling national statistics on teachers' lack of training in core content areas and the frequency of teachers working out-of-field. Most districts and states have invested in modern student information and human resource data systems to comply with NCLB reporting requirements. These new data systems made it possible for researchers to estimate individual teacher effects on student outcomes, which subsequently shaped the Obama administration's focus on teacher evaluation reform.

Annual reporting requirements under the Bush and Obama Departments of Education made states' progress toward employing highly qualified teachers and implementing teacher evaluation reforms both public and transparent. These reporting requirements influenced where state departments of education focused their attention and provided the federal government with a means of comparing states' progress. Under Obama, ED also expanded the Civil Rights Data Collection to include new measures, such as the percent of teachers in a district who are absent for at least ten days. Simply documenting that more than one in four teachers meet this definition of being chronically absent has generated considerable public pressure on schools to address teacher attendance.<sup>43</sup> The challenge is how to leverage these data and reporting requirements to advance reform efforts in a way that focuses on the underlying challenges rather than the metrics themselves.

### **Lesson 2: Competitive Grants Are Effective at Incentivizing Reforms**

Both the Bush and Obama administrations used competitive federal grants successfully to incentivize states to undertake reforms. All but four states invested substantial time and money to submit applications to Obama's RTTT competition. States were willing to change laws and propose ambitious and politically controversial reforms for the chance to win awards that amounted to only 1 percent of annual state education budgets.<sup>44</sup> While the Great Recession amplified states' willingness to compete for federal grant

dollars, such competitions can clearly motivate and provide cover for states to pursue ambitious reforms.

Competitive grant programs have several advantages. They are viewed favorably by politicians across the aisle given their voluntary nature. This is evident in the longevity of the TIF grant program, established under Bush, expanded under Obama, and continued under Trump as the Teacher and School Leader Incentive Fund. The voluntary nature of these grant programs also limits the effectiveness of those opposed to the reforms. The Obama administration effectively sidestepped initial political opposition to RTTT because teacher unions and legislators had few ways of organizing against it. However, failing to engage in interest-based discussions with this opposition proved shortsighted for the long-term success of evaluation reforms. Competitive grants are capable of catalyzing education officials to undertake major reforms, but more limited in their ability to support successful implementation.

### **Lesson 3: Reforms Have to Balance Unavoidable Conflict with the Importance of Buy-in**

History has taught us that the success of education policy initiatives depends on the will and capacity of local “street-level bureaucrats” to implement reforms.<sup>45</sup> ED under Bush and Obama invested limited time and effort soliciting educators’ diverse perspectives on their policy reform initiatives. In the case of licensure reforms, the Bush administration benefited from a less ambitious goal that enjoyed near-unanimous public support. Public support alone, however, was not enough to secure successful implementation. Facing teacher shortages in hard-to-staff regions and subject areas as well as pressure from teacher unions, state officials ultimately set low bars and wide loopholes for current teachers to demonstrate subject competency. Imposing new content-standard requirements on veteran teachers was unlikely to succeed without also providing considerable support to teachers to meet these new standards.

The evaluation reforms advanced under the Obama administration faced even stronger opposition from educators. In 2010, a national poll by Education Next found that only 24 percent of teachers were in favor of using student growth measures to inform tenure decisions.<sup>46</sup> The irony of the RTTT grant competition was that the rubric included points for evaluation systems that “are designed and developed with teacher and principal involvement,” while the prescriptive scoring rubric and short timeline simultaneously undercut

these very efforts. States could compel districts to implement new systems, but they had no control over the scores principals gave teachers. Principals overwhelmingly assigned observation ratings that ensured all but a small percentage of teachers would receive an overall rating of proficient or higher under the new evaluation systems.<sup>47</sup>

It is hard to imagine ambitious reforms benefiting from the support of all the diverse constituents in the education sector. The lesson here is not to restrict reform efforts to popular approaches or watered-down changes at the margin, but to recognize that efforts to change the status quo are likely to be more successful when reformers involve diverse stakeholders early on in the design process. Federal teacher policy reforms suggest this is particularly true when reforms require changes to educational practices instead of technical or regulatory policy changes.

#### **Lesson 4: Education Reforms Need to Be Responsive to Basic Labor Market Principles**

The teacher policy reforms under Bush and Obama resulted in a number of unintended but largely predictable consequences given basic economic principles. We know that labor markets differ substantially across local contexts. We know that raising the barriers and costs to entry into a profession will affect who chooses to pursue that profession. We know that nonpecuniary benefits such as working conditions and job security are important elements of a profession's overall compensation. Federal efforts to promote teacher licensure and evaluation reforms did little to anticipate and address the potential negative consequences of these reforms.

Uniform licensure standards set the bar too low in many local labor markets but too high in others. Requiring teachers to become certified in every subject they teach proved to be a substantial challenge for rural districts that often rely on teachers to teach multiple subjects. Hard-to-staff schools in disadvantaged neighborhoods faced similar challenges attracting and retaining highly qualified teachers. Licensure exams created a new barrier to entry into the teaching profession that was related only weakly to teacher effectiveness. Introducing high-stakes teacher evaluations made fewer people pursue teaching as a career, likely because of the perceived loss of job security and professional autonomy for teachers.

Many of the negative consequences of federal teacher policy reforms could have been addressed directly through alternative policy designs or efforts to mitigate these negative effects. States might have been given the opportunity

to offer districts temporary waivers from certain HQT requirements. Funds might have been made available for test-prep services and tutoring supports to help teachers pass licensure exams and take coursework in their subject area. States might have offset the introduction of high-stakes evaluation with new approaches to teacher compensation and investments in teachers' working conditions.

### **Lesson 5: Prioritize Clear Goals but Allow for Flexibility in Processes and Timelines**

The goals of the Bush and Obama administrations were quite sensible: strengthening teachers' content knowledge and making human capital decisions based on job performance rather than seniority. These federal education reform efforts ran aground in many states because of their prescribed set of system structures and rapid, politically advantageous implementation timeline. Instead, ED under Bush and Obama might have pursued a more flexible approach that recognized states' different starting places and levels of internal capacity.

When RTTT began in 2009, there was not a single statewide evaluation system that employed the design features rewarded by the grant rubric. The administration's consistent pressure on states not to delay their reform commitments left little time for pilot testing and professional training. States that could not meet these ambitious timelines were threatened with the loss of grant funds or waiver approvals rather than supported in adapting their approaches and timelines. The high priority placed on test-score-based measures of teacher performance created major opposition to performance-based decision making in public schools. At the same time, the wide flexibility provided to states under the HOUSSE rules allowed them to make the HQT requirements under NCLB largely a paper compliance exercise for teachers already in the classroom. Striking the right balance between flexibility and follow-through is difficult, but critical.

## **CONCLUSION**

### **The Federal Role for Improving Teacher Quality**

Starting with the passage of ESEA in 1965, the federal government has had an important and growing role in shaping education policy. The successes and failures of federal efforts to strengthen teacher quality highlighted in this chapter do not suggest that a diminished role is more appropriate. It



seems unlikely that a majority of states would have been more successful at improving the quality of the public teacher workforce or addressing the inequitable access to effective teachers in the absence of these federal efforts. What is unclear is whether the positive effects of these reform efforts could have been realized with fewer negative consequences if these administrations had taken alternative approaches. Lessons from these efforts illuminate ways in which the federal government can continue to leverage its position to improve educator effectiveness within the context of our decentralized education system.

Perhaps the most enduring effect of teacher reform efforts under Bush and Obama is the now widespread recognition of the central importance of teacher quality. This shift in perceptions has helped shape teacher policy in productive ways, such as reducing the use of seniority-based layoffs and transfer rights for tenured teachers to “bump” provisional teachers from their positions. District leaders and school administrators now have far more meaningful information about teachers’ qualifications and performance relative to when Bush or Obama first took office.

Thirty-five years after *A Nation at Risk* was first released, the challenge of improving teacher quality remains as relevant now as ever before. Progress will likely be slow and uneven, but investing in teacher quality is among the most important commitments we can make as a nation.

*Acknowledgments: I am grateful to Chad Aldeman, David Dunn, Rick Hess, Michael McShane, Mike Petrilli, Kate Walsh, Daniel Weisberg, Joanne Weiss, and Ross Wiener for the helpful guidance and insights they provided. Brendan Bell, Sarah Conlisk, and Anna Meyer provided outstanding research and editorial assistance. All errors and omissions are my own.*

