
THE PROMISES AND PITFALLS OF TEACHER EVALUATION
AND ACCOUNTABILITY REFORM*

Benjamin Michael Superfine**

* Publication forthcoming in: *The Enduring Legacy of Rodriguez: Creating New Pathways to Equal Educational Opportunity* (tentative title).

** Associate Professor in the College of Education, University of Illinois at Chicago.

I. INTRODUCTION

In the late 2000s and early 2010s, a wave of state-level laws aimed at increasing teacher effectiveness emerged across the United States. These laws generally increased the frequency of teacher evaluations, specified in detail the ways in which teachers must be evaluated, and increased accountability for teachers' performance. In doing so, these laws formally tied teacher evaluation to student performance on standardized tests and grounded major decisions about teacher tenure and employment in these evaluations. From 2009 to 2012, as many as 36 states and Washington D.C. made such changes, and from 2011 to 2012, at least 15 states modified their processes specifically governing teacher evaluation and tenure.¹ While states have been at the forefront of these changes, the federal government also has pushed states in this direction. The American Recovery and Reinvestment Act of 2009 ("ARRA") contained financial incentives for states to engage in such reform,² and waivers releasing states from certain accountability requirements of the No Child Left Behind Act of 2001 ("NCLB") allowed states to make such changes as well.³ As such, there has been unprecedented legal and political momentum to increase teacher effectiveness primarily through the mechanisms of evaluation and accountability.

This movement to increase teacher effectiveness has strong roots extending to efforts focused on equalizing students' educational opportunities that began with the struggle to desegregate schools and continued with school finance reform litigation. Since the late 1960s and early 1970s, courts and legislatures at both the federal and state levels have intensely focused on the financial resources devoted to public education.⁴ In addition to prevalent educational statistics, such as student achievement on standardized tests,⁵ school funding has constituted one of the most

¹ NAT'L COUNCIL ON TCHR QUALITY, STATE OF THE STATES 2012: TEACHER EFFECTIVENESS POLICIES 1 (2012), http://www.nctq.org/dmsView/State_of_the_States_2012_Teacher_Effectiveness_Policies_N-CTQ_Report.

² See generally American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115.

³ Alyson Klein, "*Race to the Top*" Standards Link Questioned, EDUC. WEEK, Dec. 16, 2009, at 13, available at http://www.edweek.org/ew/articles/2009/12/16/15standards_ep-2.h29.html.

⁴ Benjamin M. Superfine, New Directions in School Funding and Governance: Moving from Politics to Evidence, 98 KY. L.J. 4, 653, 653. (2009-2010).

⁵ Lynn Olson, NAEP Gains are Exclusive in Key Areas, EDUC. WEEK, Oct. 26, 2005, at 1, available at <http://www.edweek.org/ew/articles/2005/10/26/09naep.h25.html>.

important indicators to observers for gauging school quality and educational opportunity.⁶ Unfortunately, the tremendous efforts devoted to improving the distribution of educational funds have largely failed to affect students' educational opportunities in the intended fashion. Several efforts at using the courts to equalize and augment educational funding have not been successful,⁷ and even where such efforts have succeeded, it is far from clear that school quality and learning opportunities have consistently been equalized or increased.⁸

As such, reformers have turned their attention to improving other educational “inputs”, as well as focusing on the governance and “outputs” of schooling. For example, since the 1960s reformers have emphasized a range of reform strategies hinging on concepts such as curriculum, teacher quality, standards, testing and accountability, and school choice. Since the 1990s, major education reforms have focused on accountability for educational outputs and school choice, as reflected in laws such as NCLB. The move toward educational outputs reflects a political landscape much changed from the early days of desegregation and school finance reform. In an era of quickly changing technology, globalization, and the dominance of business management principles, increasing student outputs to support the international economic competitiveness of the U.S. has overtaken some historical purposes of education like promoting equal educational opportunity.⁹ Still, educational inputs continue to receive some attention as critical leverage points for improving schooling. Reforms focused on teacher evaluation and accountability reflect a fusion of approaches focused on both educational inputs and outputs.¹⁰ These reforms concentrate on improving the input perhaps closest to students – teachers – and rely heavily on the most dominant educational output currently used – student achievement scores on standardized tests.¹¹ As such, the policymaking community consistently has framed these reforms as solutions to the problems of both educational equity and the economy.¹²

Despite its recent emergence as a central element of federal and state education reform efforts, the teacher evaluation and accountability

⁶ See Michael Heise, *Litigated Learning and the Limits of Law*, 57 VAND. L. REV. 2417, 2417 (2004).

⁷ See Bradley W. Joondeph, *The Good, the Bad, and the Ugly: An Empirical Analysis of Litigation Prompted School Finance Reform*, 35 SANTA CLARA L. REV. 763, 769 (1995).

⁸ See generally David K. Cohen, et al., *Resources, Instruction, and Research*, 25 EDUC. EVALUATION & POL'Y ANALYSIS 2, 119, 119–42 (2003).

⁹ See BENJAMIN M. SUPERFINE, *EQUALITY IN EDUCATION LAW AND POLICY, 1954-2010*, at 11 (2013).

¹⁰ See *id.* at 3.

¹¹ See *id.* at 3, 211.

¹² See Stephen Sawchuk, *Stimulus Bill Spurs Focus on Teachers: Language on Fair Distribution, Effectiveness Offers Policy Clues*, EDUC. WEEK, March 11, 2009, at 1, available at <http://www.edweek.org/ew/articles/2009/03/11/24stimteach.h28.html>.

movement entails both significant promise and potential pitfalls. A growing consensus among education researchers has emerged that teacher quality is one of the most important influences on student learning and achievement.¹³ One study estimated that teacher quality explains 7.5 percent of the variation in student achievement and that teacher quality may explain up to 20 percent.¹⁴ It also appears that teacher quality is especially important for improving the performance of poor and minority students.¹⁵ However, teacher quality is inequitably distributed among students. The U.S. Department of Education found that only 23 percent of all teachers and 14 percent of teachers in high-poverty schools come from the top third of college graduates, and 90 percent of high-minority districts face significant challenges attracting highly qualified science and mathematics teachers.¹⁶ Compounding such problems, most traditional teacher evaluation systems have resulted in almost every teacher receiving a high rating. For example, the New Teacher Project found in a study of 12 school districts that 99 percent of teachers were rated as satisfactory.¹⁷ The new wave of teacher evaluation and accountability reforms accordingly represents an attempt to improve a major factor affecting the inequitable distribution (and to some, unacceptably low level) of students' learning opportunities.¹⁸

On the other hand, the laws aimed at enhancing teacher evaluation and accountability entail several potential pitfalls. The recent teacher evaluation and accountability reforms involve a host of psychometric issues that weaken the validity of determinations about whether teachers are performing acceptably.¹⁹ Even if these systems are "good enough" at evaluating teacher effectiveness, the incentives contained in these systems are likely insufficient to enhance teacher effectiveness through increasing

¹³ See Brian Rowan et al., *What Large-Scale, Survey Research Tells Us About Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools*, 104 TCHR.'s C. REC. 1525, 1537 (2002), available at <https://www.tcrecord.org/library/pdf.asp?ContentId=11041>.

¹⁴ See Eric A. Hanushek et al., *Teachers, Schools, and Academic Achievement* 32 (Nat'l Bureau of Econ. Research, Working Paper No. 6691, 1998).

¹⁵ See LAURA GOE, NAT'L COMPREHENSIVE CTR. FOR TCHR QUALITY, *THE LINK BETWEEN TEACHER QUALITY AND STUDENT OUTCOMES: A RESEARCH SYNTHESIS*, 1 (2007), available at <http://files.eric.ed.gov/fulltext/ED521219.pdf>.

¹⁶ U.S. DEP'T OF EDUC., *OUR FUTURE, OUR TEACHERS: THE OBAMA ADMINISTRATION'S PLAN FOR TEACHER EDUCATION REFORM AND IMPROVEMENT* 5–6 (2011), available at <http://www.ed.gov/sites/default/files/our-future-our-teachers-accessible.pdf>.

¹⁷ DANIEL WEISBERG ET AL., *THE NEW TEACHER PROJECT, THE WIDGET EFFECT: OUR NATIONAL FAILURE TO ACKNOWLEDGE AND ACT ON DIFFERENCES IN TEACHER EFFECTIVENESS* 6 (2d ed. 2009), available at <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.

¹⁸ See generally ORG. FOR ECON. COOPERATION AND DEV., *STRONG PERFORMERS & SUCCESSFUL REFORMERS IN EDUCATION: LESSONS FROM PISA FOR THE UNITED STATES* 32–38, 42, 49–52, (2011), available at <http://www.oecd.org/pisa/46623978.pdf>.

¹⁹ See generally STEVEN GLAZERMAN ET AL., *BROWN CTR. ON EDUC. POL'Y AT BROOKINGS, PASSING MUSTER: EVALUATING TEACHER EVALUATION SYSTEMS* 1–4 (2011).

teachers' motivation as theorized.²⁰ On a broader level, these systems raise serious problems of policy incoherence. They narrowly focus on evaluation and accountability, and are not comprehensively or strategically designed with other functions of the teacher workforce in mind (e.g. teacher recruitment, promotion, professional development).²¹ Laws focused on teacher evaluation and accountability accordingly continue to raise one of the most significant problems historically facing education policies.²² Finally, these laws reflect an underlying conception of teaching that paradoxically may hinder teacher effectiveness. They rely on bureaucratic controls associated with unskilled labor to restrict teachers' autonomy and discretion, which can be critical for helping teachers meet their goals as effectively as possible.²³ All these pitfalls are particularly problematic because they are interrelated. The narrow focus of the policies and potential problems stemming from the evaluation and accountability mechanisms flow directly from the conception of teaching undergirding these systems. Indeed, laws governing teacher evaluation and accountability have been enacted on the basis of very limited evidence that these systems actually can improve teacher effectiveness.

This chapter examines the recent wave of laws aimed at enhancing teacher evaluation and accountability, and recommends strategies for moving forward in a way that holds greater promise for providing students with more equal and greater educational opportunities. First, this chapter provides a historical overview of the legal landscape governing the primary functions of the teacher workforce. Second, the major characteristics of this new wave of laws are examined. Third, the strengths and potential pitfalls entailed by these laws are analyzed in light of educational research. Finally, this chapter offers recommendations for improving teacher evaluation and accountability laws in a way that offers significantly more promise for equalizing and improving students' educational opportunities.

II. THE LEGAL LANDSCAPE OF THE TEACHER WORKFORCE

State and local governments traditionally have held the primary authority

²⁰ See Thomas B. Corcoran, *The Changing and Chaotic World of Teacher Policy*, in *THE STATE OF EDUCATION POLICY RESEARCH* 307, 318 (Susan H. Fuhrman et al., eds., 2007).

²¹ See generally Benjamin M. Superfine et al., *The Challenging Road to Coherence in Illinois Education Policy* 73–80, in *THE ILLINOIS REPORT 2012*, INST. GOV'T & PUB. AFFAIRS (2012).

²² See David K. Cohen, *Policy and Organization: The Impact of State and Federal Education Policy on School Governance*, 52 *HARV. EDUC. REV.* 474, 481 (1982).

²³ See generally Richard Pratte & John L. Rury, *Teachers, Professionalism, and Craft*, 93 *TCHR.'S C. REC.* 59, 64–67 (1991).

to govern education and teachers in particular.²⁴ While the legal authority to govern teachers generally has resided at the state level, a range of decisions about teacher management traditionally has been made at the district and school level.²⁵ Teachers have long been considered among the most important actors in the education policy process, and school level administrators historically have been the primary entities responsible for selecting teachers' curricula and evaluating teachers.²⁶ As school districts grew and federal and state governments enacted laws governing collective bargaining in the first half of the 20th century, a substantial amount of authority became situated at the school district level.²⁷ Through collective bargaining processes, most school districts and teachers' unions now agree on teacher compensation and procedural protections for teachers' employment rights (within the boundaries set by state law).²⁸ Such processes have often generated complex and often cumbersome processes for teacher evaluation and dismissal.²⁹

Still, states have taken on increasingly important roles in education and teacher governance throughout the 20th century. States' legal bases for governing education and teacher policy generally stem from state constitutions. Although the U.S. Constitution does not mention education, every state constitution specifies at least some legal duty that the state must provide students with a system of public schools.³⁰ All states have laws requiring teachers to hold licenses.³¹ However, the nature and quality of licensure standards and assessments vary widely. These laws vary with regard to the college courses that must be taken, clinical experiences, and scores on assessments measuring general knowledge, subject matter knowledge, and pedagogical knowledge.³² States also possess the legal

²⁴ See generally Cohen, *supra* note 22, at 476–82.

²⁵ See generally LARRY CUBAN, *HOW TEACHERS TAUGHT: CONSTANCY AND CHANGE IN AMERICAN CLASSROOMS 1890-1980*, at 53–55 (1984) (“One of the strongest signals to teachers on what they [are] to do in class is the evaluation rating and the manner that supervision [is] implemented by principals, first assistants, district superintendents, and department chairmen.”).

²⁶ See generally *id.* at 264 (“Educational policymakers have generally given a rhetorical bow to the importance of the classroom teacher and school principal in implementing instructional or curricular change and then assigned to each organizational level what tasks were to be done.”).

²⁷ See Paul T. Hill, *The Costs of Collective Bargaining Agreements and Related District Policies*, in *COLLECTIVE BARGAINING IN EDUCATION: NEGOTIATING CHANGE IN TODAY'S SCHOOLS* 89, 89 (Jane Hannaway & Andrew J. Rotherham, eds., 2006).

²⁸ See *id.* at 102–07.

²⁹ See generally *id.* at 99 (“The result [of collective bargaining agreements] . . . is that principals must often oversee teaching staffs that they do not hire, cannot fire, and that are perfectly free to cooperate with or oppose any effort to improve instruction.”).

³⁰ Peter Enrich, *Leaving Equality Behind: New Directions in School Finance Reform*, 48 *VAND. L. REV.* 101, 106 (1995).

³¹ Corcoran, *supra* note 20, at 309.

³² *Id.*

authority to approve teacher education and training programs. However, the requirements for these programs also vary with regard to whether programs are approved by the state, must meet national accreditation standards, or must meet state standards. States similarly have enacted laws governing professional development requirements for teachers, but these requirements also vary significantly.³³

State laws have long set teacher compensation requirements as well. These laws generally dictate teacher pay according to state or local salary schedules on the basis of college credits, experience, and credentials.³⁴ Some states have modified teacher evaluation laws through policies such as those that are put in place by career ladders for teachers and merit pay systems.³⁵ However, more recent reforms also vary significantly and include approaches such as signing bonuses, housing supplements, higher pay in shortage areas (such as mathematics), higher pay for teachers in hard-to-staff schools, and skill-based incentives.³⁶ Although states have found it difficult to fully implement such alternative teacher compensation policies,³⁷ their increasing traction reflects the strong political momentum for using state law to increase teacher effectiveness.

The standards-based reform and accountability movements marked an important, though indirect, increase in the legal authority of both state and federal governments over teacher policy. Emerging in states in the mid-1980s, the standards-based reform movement focused on both “bottom up” and “top down” change. Under this movement, states enacted standards articulating what a student should know and be able to do after receiving instruction from any teacher.³⁸ Yet, state standards largely left decisions about how to teach in teachers’ hands.³⁹ In 1994, the federal government became formally involved in this movement with the passage of the Goals 2000: Educate America Act (Goals 2000) and the Improving America’s School Act (“IASA”).⁴⁰ While Goals 2000 provided funding to states to

³³ Table Comparing Professional Development Requirements by State, EDUCATION COUNTS, <http://www.edweek.org> (follow “Education Counts Database” hyperlink; then select “Custom Table Builder”; then select “Teaching profession” folder; then select “Professional development” folder; then check boxes for indicators; then select “2012”; then select “Select All” for states; then select “View Table”) [hereinafter EdCounts Database].

³⁴ Corcoran, *supra* note 20, at 316.

³⁵ *Id.* at 317.

³⁶ *Id.*

³⁷ *Id.* at 320.

³⁸ Paul E. Barton, The Elementary and Secondary Education Act and Standards-Based Reform, in *THE FUTURE OF THE FEDERAL ROLE IN ELEMENTARY & SECONDARY EDUCATION* 61, 61 (2001)

³⁹ See *id.* at 63–66.

⁴⁰ David K. Cohen & Susan L. Moffitt, Title I: Politics, Poverty, and Knowledge, in *THE FUTURE OF THE FEDERAL ROLE IN ELEMENTARY & SECONDARY EDUCATION* 77, 77 (2001); see generally Improving America’s School’s Act of 1994, Pub. L. No. 103-382, 108 Stat. 3518 (1994).

facilitate the development of their own standards and assessment systems,⁴¹ the IASA conditioned the continuing receipt of Title I funds from the Elementary and Secondary Education Act on the development of these systems.⁴² This emphasis on educational outcomes and accountability put additional pressure on teachers to teach particular content and use particular methods to achieve student mastery of skills and knowledge specified by state standards.⁴³ Moreover, a multi-state consortium is currently developing the Common Core State Standards – student learning standards in mathematics and language arts that are being adopted by 45 states and the District of Columbia.⁴⁴ Federal and state laws accordingly have become important drivers of teachers’ instructional work.⁴⁵

With the enactment of NCLB in early 2002, the role of the federal government in teacher policy dramatically increased.⁴⁶ NCLB required states to develop and implement standards and testing systems, and hold schools accountable for student performance on standardized tests.⁴⁷ Building on the requirements of many existing state laws, Goals 2000, and the IASA, NCLB resulted in significant pressure on many teachers to “teach to the test” by engaging in practices such as showing students how to fill in answer sheets and emphasizing test-taking skills in lieu of more “meaningful” learning experiences.⁴⁸ NCLB also focused directly on improving teacher effectiveness by requiring states to place a highly qualified teacher (“HQT”) in every public school classroom.⁴⁹ While states were granted some flexibility under the law, NCLB generally defined an HQT as having a bachelor’s degree, being fully certified, and having demonstrated knowledge and skills in his or her field.⁵⁰ Since the passage of NCLB, the HQT requirements have been strongly criticized on a variety of

⁴¹ Cohen & Moffitt, *supra* note 40, at 82.

⁴² See Improving America’s Schools Act of 1994, Pub. L. No. 103-382, tit. I, § 1111(a)(1) & (b)(1), 108 Stat. 3965 (1994) (codified as amended at 20 U.S.C. § 6311(a)(1) & (b)(1) (2006)).

⁴³ See *id.* at tit. I, § 1001 (d)(1) – (8).

⁴⁴ See The Standards, Common Core State Standards Initiative, <http://www.corestandards.org/the-standards> (last visited Sept. 26, 2013) (showing the standards apply to mathematics and language arts); see also In the States, Common Core State Standards Initiative, <http://www.corestandards.org/in-the-states> (last visited Sept. 26, 2013) (showing states that have adopted the Common Core Standards).

⁴⁵ See generally Improving America’s Schools Act of 1994 (defining how state plans for academic standards must fit within the federal scheme).

⁴⁶ See No Child Left Behind Act of 2001, 20 U.S.C. § 6301 (2006).

⁴⁷ *Id.* at § 6316(a).

⁴⁸ CENTER ON EDUC. POL’Y, FROM THE CAPITAL TO THE CLASSROOM: YEAR 4 OF THE NO CHILD LEFT BEHIND ACT 103 (March 2006), available at <http://www.cep-dc.org/displayDocument.cfm?DocumentID=301>.

⁴⁹ No Child Left Behind Act of 2001 § 6319(a)(2).

⁵⁰ Archived Info., U.S. Dept. of Educ., No Child Left Behind: A Toolkit for Teachers, Highly Qualified Teacher Requirements (last visited Sept. 29, 2013), http://www2.ed.gov/teachers/nclbguide/toolkit_pg10.html.

grounds, most notably for being too weak to increase teacher effectiveness and failing to address the fundamental problems of teacher performance that are not captured by indicators such as degrees held.⁵¹

While a broad overview of the legal landscape governing teachers generally reflects the growth of requirements specifying what teachers should do in the classroom and an increasing focus on teacher quality and effectiveness, the trajectory of approaches toward teaching has not been completely even. For example, competency tests of teachers' basic skills were administered to prospective teachers in most states in the beginning of the 20th century, but the use of these tests declined as states enacted requirements for teachers to graduate college and be certified.⁵² During the early 1980s, many states participated in the "educational excellence" movement by increasing student testing and the use of standardized curricula.⁵³ However, especially in light of criticism that such an approach "deskilled" teachers by weakening their ability to exercise their discretion, a second wave of state-level reform that emerged in the mid-1980s aimed at restructuring schools to promote school-level autonomy and flexibility.⁵⁴ Indeed, articulating particular learning objectives and student outcomes has constituted a key approach employed by state and federal governments to improve teacher effectiveness.⁵⁵ Yet, such an approach has generated strong concerns that these reforms paradoxically decrease teacher effectiveness in the end.⁵⁶

A. Recent Federal and State Teacher Evaluation and Accountability Reforms

Since 2009, federal and state governments have become deeply involved in attempts to increase teacher effectiveness through the widespread passage of laws governing teacher evaluation and accountability. Aimed at stabilizing and stimulating the U.S. economy during its worst financial crisis since the Great Depression, the ARRA included several provisions

⁵¹ BARNETT BERRY ET AL., CENTER FOR TEACHING QUALITY, NO CHILD LEFT BEHIND AND THE 'HIGHLY QUALIFIED' TEACHER: THE PROMISE AND POSSIBILITIES 1-2 (2006).

⁵² Martha M. McCarthy, Teacher-Testing Programs, in THE EDUCATIONAL REFORM MOVEMENT OF THE 1980s 189, 189 (Joseph Murphy ed., 1990).

⁵³ William L. Boyd, Balancing Control and Autonomy in School Reform: The Politics of Perestroika, in THE EDUCATION REFORM MOVEMENT OF THE 1980s 85, 86 (Joseph Murphy ed., 1990).

⁵⁴ Id. at 86-87.

⁵⁵ See LINDA M. MCNEIL, CONTRADICTIONS OF CONTROL: SCHOOL STRUCTURE AND SCHOOL KNOWLEDGE 161-62 (1986).

⁵⁶ See id. at 162.

directly aimed at increasing teacher effectiveness.⁵⁷ Under this law, states were required to institute data systems linking student and teacher performance data and teacher incentive policies in exchange for receiving funds to minimize educational reductions. The Race to the Top Fund (“RTT”) contained in the ARRA, a \$4.35 billion competitive grant program targeted at spurring innovation in education, also provided grants to winning states partly to implement policies focused on tying decisions about educator performance to student achievement data.⁵⁸

The federal scoring rubric contained in federal regulations for states’ RTT applications reflected this emphasis.⁵⁹ The rubric awarded points to state’s RTT applications based on a range of factors, including states’ commitment to improving standards and assessments, building more robust data systems to support instruction, and improving teacher effectiveness based on performance.⁶⁰ Taking a less formal approach, U.S. Secretary of Education Arne Duncan made several public statements pressing states to enact reforms focused on increasing teacher evaluation and accountability. For example, he singled out California, New York, and Wisconsin as committing “a grave disservice to the teaching profession and to our nation’s children” because they maintained a “firewall” between student and teacher data.⁶¹

Although the ARRA constituted the most high-profile effort focused on using teacher evaluation and accountability to improve teacher effectiveness, several pre-existing reforms existed at the state level as well. For example, Tennessee long had in place a database linking growth in student achievement on standardized test to individual teacher, and since at least 2000 required teachers to be evaluated in six areas (such as planning and assessment of students).⁶² Other states, such as Colorado, had begun to engage in similarly oriented efforts before the passage of the ARRA by allowing some districts to experiment with merit pay for strong evaluations of teachers’ performance.⁶³ Indeed, as early as 2003, nine states had in place policies encouraging teachers to be paid for performance, although only two

⁵⁷ American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115, 283.

⁵⁸ U.S. DEPT. EDUC., RACE TO THE TOP EXECUTIVE SUMMARY 2 (2009), available at <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>.

⁵⁹ Race to the Top Fund, 74 Fed. Reg. 59,473, 59,813 (Nov. 18, 2009).

⁶⁰ *Id.* at 59,836.

⁶¹ Michele McNeil & Stephen Sawchuk, Rich Prize, Restrictive Guidelines, 28 EDUC. WEEK 1, 2 (May 15, 2013), available at <http://www.edweek.org/ew/articles/2009/08/07/37stimrace.h28.html>.

⁶² J.E. Stone, *Policy Highlights from Tennessee’s Race to the Top Application*, EDUC. CONSUMERS FOUND., Feb. 15, 2010, at 7.

⁶³ Michele McNeil, Racing for an Early Edge, 28 EDUC. WEEK 1, 5 (May 15, 2013), available at http://www.edweek.org/ew/articles/2009/07/09/36stimulus_ep.h28.html.

states had in place policies that tied teacher evaluation to student achievement.⁶⁴

Given states' budget woes, the lure of ARRA funds, and federal pressure, several other states began to enact laws focused on teacher evaluation and accountability since 2009. Each of the 14 states that won RTT grants promised to use student achievement data as a significant part of teacher evaluation and teacher evaluation data to make personnel decisions.⁶⁵ However, the push for increased teacher evaluation and accountability has spread beyond the scope of RTT funding distribution. As reported by the National Council on Teacher Quality, only 14 states required annual evaluations of all teachers in 2009, while 23 states required annual evaluations for all teachers and 43 states required annual evaluations of all new teachers in 2012.⁶⁶ Only 4 states required student achievement to be an important factor in assessing teacher performance in 2009, and 20 states required student achievement to be at least a significant factor for judging teacher performance in 2012.⁶⁷ In the same period, the number of states requiring teacher evaluations to include the collection of data of student learning jumped from 15 to 30.⁶⁸ And while 17 states required teacher ratings to be broken down into multiple performance levels in 2011, 25 states required teacher evaluation systems to include multiple levels in 2012.⁶⁹ To be sure, there are several potential approaches for states to link student and teacher data. However, states generally have focused on value-added modeling ("VAM"), a process aimed at linking students' test score growth over a period of time, such as a school year, to an individual teacher.⁷⁰

Although student achievement is a central component of many teacher evaluation systems, these systems include several other components as well. Many laws require evaluation systems to incorporate student ratings of teachers and direct assessments of teacher knowledge in addition to student test score gains in teachers' classrooms.⁷¹ Many of these systems also include classroom observations conducted by administrators, who are primarily the teachers' principals.⁷² In 2012, 39 states required annual

⁶⁴ EdCounts Database, *supra* note 33.

⁶⁵ Benjamin M. Superfine et al., *The Expanding Federal Role in Teacher Workforce Policy*, 26 *EDUC. POL'Y*, 58, 64 (2012).

⁶⁶ NAT'L COUNCIL ON TCHR QUALITY, *supra* note 1, at 1.

⁶⁷ *Id.* at 5.

⁶⁸ *Id.* at 1.

⁶⁹ *Id.* at 2.

⁷⁰ *Id.* at 6.

⁷¹ GLAZERMAN ET AL., *supra* note 19 at 3.

⁷² *Id.* at 3, 14.

classroom observations of teachers.⁷³ Some of these systems have included other sources of information as well, including ratings of teachers' commitment to the school community, measures of teacher professionalism (such as unexcused teacher absences and late arrivals), and student test score gains for a teacher's school.⁷⁴ States' teacher evaluation systems notably differ with regard to the autonomy accorded to school districts to develop their own methods of evaluation. For example, Rhode Island provides districts with the authority to develop their own evaluation systems, so long as they adhere to certain broad requirements. These requirements include evaluating teachers annually and rating teachers as falling into one of four categories ranging from "Ineffective" to "Highly Effective."⁷⁵ In contrast, North Carolina centrally determines the components, measures, and frequency of evaluations, and generally requires districts to implement this model with little discretion.⁷⁶

B. Recent Teacher Preparation Program Reforms

The push for evaluating teachers and holding them accountable has begun to bleed into state laws governing teacher preparation as well. States traditionally have acted as the primary regulators of the quality of teacher preparation programs.⁷⁷ State educational agencies generally review teacher preparation programs to ensure that they meet state law and regulations, and they approve programs on the bases of these reviews.⁷⁸ In addition, states set the requirements for teacher certification, which generally include passing content and pedagogical tests and completing an approved program.⁷⁹ Several states additionally require or encourage teacher preparation programs to be accredited by independent agencies like the

⁷³ NAT'L COUNCIL ON TEACHER QUALITY, *supra* note 1, at 2.

⁷⁴ GLAZERMAN ET AL., *supra* note 19, at 3.

⁷⁵ R.I. DEP'T OF EDUC., EDUCATOR EVALUATION SYSTEM STANDARDS (2009), available at <http://www.ride.ri.gov/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/Educator-Evaluation/Ed-Eval-Standards/EdEvalStandards.pdf>.

⁷⁶ CTR. ON GREAT TCHRS. & LEADERS AT AM. INSTS. FOR RESEARCH, Databases on State Teacher and Principal Evaluation Policies, <http://resource.tqsource.org/stateevaldb/StateRoles.aspx> (last accessed Sept. 25, 2013).

⁷⁷ See JANE G. COGGSHALL ET. AL., NAT'L COMPREHENSIVE CTR. FOR TCHR. QUALITY, EVALUATING THE EFFECTIVENESS OF TEACHER PREPARATION PROGRAMS FOR SUPPORT AND ACCOUNTABILITY 2 (2012), available at http://www.gtlcenter.org/sites/default/files/docs/TQ_RandP_BriefEvaluatingEffectiveness.pdf.

⁷⁸ *Id.*

⁷⁹ *Id.*

National Council for the Accreditation of Teacher Education or the Teacher Education Accreditation Council.⁸⁰

Despite the presence of these quality control mechanisms, strong attacks have been levied at teacher preparation programs from several directions in recent years. For example, in 2008 the National Council on Teacher Quality found that teacher preparation programs for elementary mathematics teachers use inadequate textbooks, have low standards for admission and graduation, are not rigorous, and employ teacher educators who are not sufficiently qualified to teach mathematics.⁸¹ U.S. Secretary of Education Arne Duncan stated, “[b]y almost any standard, many if not most of the nation’s 1,450 schools, colleges, and departments of education are doing a mediocre job of preparing teachers for the realities of the 21st-century classroom,” and, “America’s university-based teacher preparation programs need revolutionary change – not evolutionary tinkering.”⁸²

Given such criticism, teacher preparation policies have begun to change at federal and state levels. Under the Higher Education Act, states must report 440 different types of data annually about teacher preparation.⁸³ However, ED has argued that such data are not “meaningful indicators” of program effectiveness.⁸⁴ As such, a federal rule-making panel engaged in discussions in mid-2012 to modify federal reporting regulations to emphasize outcome measures of teacher preparation programs, including the achievement growth of students served by teachers from these programs.⁸⁵ However, the rule-making process ultimately stalled due to the controversy surrounding such changes.⁸⁶ Significant changes also have emerged at the state level. As reported by the National Comprehensive Center for Teacher Quality, at least six states have enacted significant changes to their policies governing teacher preparation programs.⁸⁷ Louisiana has begun to use the same value-added measure of student performance on achievement tests to evaluate both in-service teachers and

⁸⁰ *Id.*

⁸¹ See NAT’L COUNCIL ON TCHR. QUALITY, *NO COMMON DENOMINATOR: THE PREPARATION OF ELEMENTARY TEACHERS IN MATHEMATICS BY AMERICA’S EDUCATION SCHOOLS* 35, 38, 40, 46 (2008), available at http://www.nctq.org/p/publications/docs/nctq_ttmath_fullreport_20090603062928.pdf.

⁸² Arne Duncan, Sec’y, Dep’t of Educ., Remarks at Teachers College, Columbia University, *Teacher Preparation: Reforming the Uncertain Profession* (Oct. 22, 2009) (transcript available at <http://www2.ed.gov/news/speeches/2009/10/10222009.html>).

⁸³ U.S. DEP’T OF EDUC., *supra* note 16, at 9.

⁸⁴ *Id.*

⁸⁵ Libby Nelson, Rule making on teacher preparation programs fails to reach consensus, *INSIDE HIGHER ED*, (Apr. 13, 2012, 3:00 AM), <http://www.insidehighered.com/news/2012/04/13/rule-making-teacher-preparation-programs-fails-reach-consensus>.

⁸⁶ *Id.*

⁸⁷ COGGSHALL ET. AL., *supra* note 77, at 23.

teacher preparation programs, in addition to modifying its certification requirements and using a web-based performance assessment system to evaluate teacher candidates.⁸⁸ Texas similarly uses value-added measures of student performance, in addition to surveying principals about the quality of teachers and examining the pass rates of teacher candidates on certification exams to evaluate teacher preparation programs.⁸⁹ Moreover, at least 13 states and the District of Columbia have reported that they plan to report value-added information on their teacher preparation programs in the coming years.⁹⁰

C. Political Pushback to Teacher Evaluation and Accountability Reforms

Although policymakers and many independent groups have strongly supported teacher evaluation and accountability reform, these reforms have also generated harsh criticism. For example, Chicago Teachers' Union ("CTU") President Karen Lewis repeatedly argued during the 2012 CTU teacher strike that reforms increasing in-service teacher evaluation and accountability on the basis of student achievement deskill teachers.⁹¹ Lewis accordingly argued that the effects of laws instituting teacher evaluation and accountability systems should be minimized to the greatest extent possible. Randi Weingarten, President of the American Federation of Teachers, similarly cautioned that teacher evaluation systems are the most beneficial for students when they are primarily aimed at the continuous improvement of classroom teaching instead of annual personnel decisions.⁹²

Such criticism has influenced the passage of federal and state laws and regulations. Texas refused to follow the lead of the Obama administration in key reform areas such as standards, and along with eight other states, provided early signals that they would not apply for RTT funding.⁹³ Echoing historical concerns about the expansion of the federal role in education, Republican leaders in Congress argued that the education reform provisions of the ARRA could result in an undesirable expansion of federal

⁸⁸ Id. at 25–26.

⁸⁹ Id. at 27–28.

⁹⁰ Stephen Sawchuk, *'Value Added' Proves Beneficial to Teacher Prep.*, 21 EDUC. WEEK 1, 3 (Feb. 22, 2012), available at http://www.edweek.org/ew/articles/2012/02/17/21louisiana_ep.h31.html.

⁹¹ Monica Davey & Steven Yaccino, *Teachers End Chicago Strike on Second Try*, N.Y. TIMES, Sept. 18, 2012, <http://www.nytimes.com/2012/09/19/us/vote-scheduled-on-chicago-teachers-contract.html?pagewanted=all&pagewanted=print>.

⁹² Press Release, Randi Weingarten, President, American Federation of Teachers, *Statement on the NCTQ Report on Teacher Evaluation* (Oct. 26, 2011), available at <http://www.aft.org/newspubs/press/2011/102611.cfm>.

⁹³ McNeil, *supra* note 63, at 22–23.

power.⁹⁴ Such political pushback emerged in state regulatory processes as well. For example, a 21-person committee of principals, superintendents, and teachers in Maryland was convened after the state won RTT funding to provide advice to the state about constructing its teacher evaluation system.⁹⁵ This committee voted 13-7 in favor of regulations allowing 30% of a teacher's evaluation to be based on student achievement growth.⁹⁶ However, all seven teachers on the committee opposed the regulations and sent a letter to the governor indicating that the policy would not provide actionable feedback to teachers and principals.⁹⁷

In short, the push for enhancing teacher effectiveness has grown quickly since the late 2000s. States increasingly have enacted laws requiring the implementation of teacher evaluation and accountability systems. This momentum has begun to spread to the legal and regulatory schemes governing teacher preparation programs. Still, many critical questions about these reforms remain unanswered. Most importantly, are these reforms efficacious and effective for equalizing and increasing students' learning opportunities? The following section discusses major education research bearing upon these laws.

III. RESEARCH ON TEACHER EVALUATION AND ACCOUNTABILITY REFORMS

The recent laws focused on teacher evaluation and accountability entail significant promise and several major challenges. These reforms represent a step in the right direction by directly engaging with teachers, who are one of the most important factors influencing students' learning opportunities. However, the reforms also involve several interlocking problems, ranging from technical psychometric issues to their underlying conceptions of effective teaching. Both the strengths and weaknesses of these reforms too often have been misconstrued in political and policy debates about education. Indeed, these reforms are highly politicized because they often cut to the core of collective bargaining agreements between teacher unions and school districts – teacher evaluation and accountability reforms directly

⁹⁴ Klein, *supra* note 3, at 13–14.

⁹⁵ Md. Exec. Order No. 01.01.2010.12, 37 Md. Reg. 830 (Jun. 18, 2010).

⁹⁶ MARYLAND COUNCIL FOR EDUCATOR EFFECTIVENESS, INITIAL RECOMMENDATIONS STATEWIDE EDUCATOR EVALUATION SYSTEM 7 (Jun. 21, 2011), available at www1.pgcps.org/Work-Area/DownloadAsset.aspx?id=180884.

⁹⁷ Letter from Betty Weller et. al., Vice President, Maryland State Education Association, to Martin O'Malley, Governor, Maryland (Jun. 20, 2011), available at http://www.washingtonpost.com/blogs/answer-sheet/post/policy-makers-ignore-the-teachers--again/2011/06/24/AGTCBEjH_blog.html.

involve issues of compensation, hiring and firing, and career advancement.⁹⁸ This section examines actual evidence bearing on teacher evaluation and accountability reforms and serves as a foundation for the recommendations in the concluding section about how these reforms should move forward.

On a very broad level, the teacher evaluation and accountability reforms appear to be moving in a positive direction. These laws are generally intended to ensure that the most effective teachers are hired and retained, and that low-performing teachers are incentivized to improve. As discussed above, teachers are one of the most important factors driving student learning.⁹⁹ Yet, teachers traditionally have not been subject to meaningful evaluations on any consistent basis.¹⁰⁰ Teacher evaluation systems traditionally have failed at distinguishing between strong and weak teachers, failed at facilitating the removal of low-performing teachers and rewarding high performing teachers, and have not provided teachers with the information that they need to improve their own instructional practices.¹⁰¹ In this respect, the reliance of teacher evaluation and accountability reforms on VAM is potentially beneficial. Because VAM incorporates students' test scores when they are first assigned to individual teachers, VAM arguably forms the basis for fairer comparisons of teachers than judgments based on their students' test scores at only a single point in time.¹⁰² VAM in teacher evaluation accordingly has the potential to improve a range of administrators' human resources decisions, including teacher pay or promotion up a career ladder.¹⁰³

Teacher evaluation and accountability reforms also have the potential to improve the equity of students' learning opportunities. Given the difficulty of hiring and retaining effective teachers in poor and high-minority schools, these reforms offer a well-defined path for improving teacher effectiveness in these schools.¹⁰⁴ Indeed, by giving educator workforce development such a prominent position, both federal and state governments have altered the general politics surrounding this area, both by placing this reform on the

⁹⁸ LAUREN SARTAIN ET AL., CONSORTIUM ON CHI. SCH. RESEARCH, UNIV. OF CHI. URBAN EDUC. INST., *RETHINKING TEACHER EVALUATION: FINDINGS FROM THE FIRST YEAR OF THE EXCELLENCE IN TEACHING PROJECT IN CHICAGO PUBLIC SCHOOLS* 1 (Jun. 2010).

⁹⁹ Linda Darling-Hammond, *Teacher Quality and Student Achievement: A Review of State Policy Evidence*, 8 *EDUC. POL'Y ANALYSIS ARCHIVES* 1, 2 (2000), available at <http://epaa.asu.edu/ojs/article/view/392>.

¹⁰⁰ SARTAIN ET AL., *supra* note 98, at 1.

¹⁰¹ *Id.*

¹⁰² EVA L. BAKER ET AL., *ECON. POL'Y INST., PROBLEMS WITH THE USE OF STUDENT TEST SCORES TO EVALUATE TEACHERS* 2 (Aug. 2010), available at <http://www.epi.org/publication/bp278/>.

¹⁰³ GLAZERMAN ET AL., *supra* note 19, at 6.

¹⁰⁴ Kati Haycock & Eric Hanushek, *An Effective Teacher in Every Classroom*, 10 *EDUC. NEXT* 47, 48-49 (2010), available at <http://educationnext.org/an-effective-teacher-in-every-classroom>.

policy agenda, and by creating emphasis and defining the “rules of engagement.” Still, as the following sections discuss, teacher evaluation and accountability reforms involve several problems that weaken their potential for improving and equalizing students’ learning opportunities.

A. Psychometric Problems

Recently enacted teacher evaluation and accountability reforms raise serious psychometric problems. The Standards for Educational and Psychological Testing, developed jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, state that testing practices must be validated and have sufficient reliability for each intended use.¹⁰⁵ In other words, evidence and theory should adequately support the interpretations of test scores entailed by the proposed uses of tests. However, teacher evaluation and accountability reforms generally employ VAM, which incorporates data from student achievement tests that have not been validated or deemed reliable for these models.¹⁰⁶

For the most part, standardized tests of student achievement have not been designed specifically for teacher evaluation. They accordingly may involve “ceiling” or “floor” effects.¹⁰⁷ Because many of these tests focus on students’ mastery of grade level standards, they are often ineffective at determining scores for high or low achieving students. As such, it is inappropriate to simply incorporate existing student achievement tests into new teacher evaluation systems. For example, an analysis of the TVAAS, the teacher evaluation system employed by Kentucky, indicated that there was not sufficient evidence and theory underlying the system’s use of student achievement data.¹⁰⁸ To validate these systems adequately, states would need to conduct validity studies, which draw in part on student data from large-scale field tests and, if necessary, operational administrations of the assessments and special research studies.¹⁰⁹ This process would involve

¹⁰⁵ AMERICAN EDUCATIONAL RESEARCH ASSOCIATION ET AL., STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING 2, 9, 17 (1999).

¹⁰⁶ Audrey Amrein-Beardsley, Methodological Concerns about the Education Value-Added Assessment System, 37 EDUC. RESEARCHER 65, 66 (2008).

¹⁰⁷ RODNEY S. WHITEMAN ET AL., CTR. FOR EVALUATION & EDUC. POLICY, REVAMPING THE TEACHER EVALUATION PROCESS 7 (2011), available at http://ceep.indiana.edu/projects/PDF/PB_V9-N4_2011_EPB.pdf.

¹⁰⁸ GENE V. GLASS, ARIZONA STATE UNIVERSITY, EDUC. POL’Y RESEARCH UNIT, TEACHER EVALUATION: POLICY BRIEF 7.10–7.15 (2004), available at <http://nepc.colorado.edu/files/EPSL-0401-112-EPRU.pdf>.

¹⁰⁹ JOAN L. HERMAN ET AL., ASSESSMENT & ACCOUNTABILITY COMPREHENSIVE CTR., DEVELOPING &

convening groups that include psychometricians, expert teachers, and experts in subject matter, instruction, English language learners, students with disabilities, and culturally diverse students. However, such processes generally have not been undertaken across the states.

Given the lack of strong validity studies to support teacher evaluation and accountability reforms, VAM estimates of teacher performance have proven very inconsistent.¹¹⁰ Evaluations of teacher performance can differ from class to class, year to year, and test to test.¹¹¹ These evaluations can differ when different statistical models are used.¹¹² Flowing from problems such as floor and ceiling effects, VAM scores for teachers have proven particularly unstable at the highest and lowest ends.¹¹³ This is particularly worrisome because these scores are the ones most likely to be used to dismiss teachers who are believed to be ineffective or reward teachers who are believed to be highly effective. As reported by the Economic Policy Institute, one study of five large urban districts employing VAM for teacher evaluation found that, among teachers ranked in the top 20% one year, less than 33% were in that group the following year, and another third moved to the bottom 40%.¹¹⁴ Another study found that VAM scores for teachers in one year could only predict 4% to 16% of the variation in these ratings in the next year.¹¹⁵ This seeming misclassification of teachers can be amplified by variations in class size or if students are not randomly assigned to teachers.¹¹⁶

The difficulty of consistently rating teachers' effectiveness also appears to stem from the huge number of factors besides teachers that influence student performance. Perhaps most importantly, VAM generally does not account for the characteristics of students that can influence the progress students make, such as family income, ethnicity, ability, and the range of other influences on students outside of school.¹¹⁷ As such, teachers who teach several English language learners or students with disabilities have shown lower student achievement gains than when they teacher other students.¹¹⁸ Moreover, there are several factors inside of schools that

SELECTING ASSESSMENTS OF STUDENT GROWTH FOR USE IN TEACHER EVALUATION SYSTEMS 4 (2011), available at http://www.cse.ucla.edu/products/policy/shortTermGrowthMeasures_v6.pdf.

¹¹⁰ BAKER ET AL., *supra* note 102, at 2.

¹¹¹ *Id.* at 14.

¹¹² *Id.* at 12.

¹¹³ *Id.*

¹¹⁴ *Id.* at 2.

¹¹⁵ *Id.* at 13.

¹¹⁶ WHITEMAN ET AL., *supra* note 107, at 4.

¹¹⁷ *Id.* at 114.

¹¹⁸ LINDA DARLING-HAMMOND, STANFORD CTR. OPPORTUNITY POL'Y IN EDUC., *CREATING A COMPREHENSIVE SYSTEM FOR EVALUATING AND SUPPORTING EFFECTIVE TEACHING* iii, available at

influence the capacity of VAM to score teacher effectiveness well. In a typical school, teaching is very interdependent – if a science teacher emphasizes mathematical explorations and computation, the science teacher could influence a student’s mathematical proficiency.¹¹⁹ Yet, the students’ mathematics teacher generally would be the one evaluated and held accountable for that student’s performance on a mathematics test. Students’ previous teachers also could influence their progress, in addition to tutors or instructional specialists, quality of curriculum materials, class size, student attendance, and out-of-school learning opportunities at places like home, libraries, online, and in the community.¹²⁰ Indeed, VAM has much difficulty disentangling the various influences on student achievement gains on standardized tests, which raises worries that a range of other factors are actually measured instead of the teacher effect. Looking broadly at such evidence, the National Research Council’s Board on Testing and Assessment stated, “VAM estimates of teacher effectiveness should not be used to make operational decisions because such estimates are far too unstable to be considered fair or reliable.”¹²¹

In addition to incorporating VAM, many recently enacted teacher evaluation systems include observational components.¹²² While much more research still needs to be completed on these components and how they are integrated with VAM, there are some early reports about their implementation. For example, an early analysis of the Framework for Teaching, the observational component used in Chicago Public Schools (CPS), indicates that principals and trained experts used the rating scale consistently and many more teachers were identified as low performing under the Framework – while only 0.3% of teachers in CPS were rated as unsatisfactory in previous years, 8% of teachers received at least one unsatisfactory rating under a study on the implementation of the Framework.¹²³

However, there were some concerns with the implementation of the Framework. Compared to external evaluators, some principals (who serve

http://www.smmcta.com/docs/Evaluation%20Research_Stanford_2012.pdf.

¹¹⁹ Glass, *supra* note 108, at 7.10.

¹²⁰ DARLING-HAMMOND, *supra* note 118, at 21.

¹²¹ NATIONAL ACADEMIES, LETTER REPORT TO THE U.S. DEPARTMENT OF EDUCATION ON THE RACE TO THE TOP FUND 10 (2009).

¹²² See, e.g., AchieveNJ: Teacher Evaluation and Support 2013-14, <http://www.state.nj.us/education/AchieveNJ/intro/1PagerTeachers.pdf>; Classroom Observation Options, UNITED FEDERATION OF TEACHERS, <http://www.uft.org/our-rights/teacher-evaluation/classroom-observation-o-ptions> (last visited Sept. 30, 2012); CPS Teacher Evaluation Plan: CPS Final Proposal, CHICAGO PUBLIC SCHOOLS, http://www.ctunet.com/quest-center/research/text/cps-framework/CPS_Final_Off-er.pdf (Mar. 29, 2012).

¹²³ SARTAIN ET AL., *supra* note 98, at 4, 15.

as the primary evaluators using the observational component) were more lenient, and some were more severe.¹²⁴ Principals also found it difficult to evaluate how well teachers communicated with students, used assessment in instruction, and organized physical space.¹²⁵ And principals generally rated teachers higher on student engagement than external observers, which is considered the most important component of the Framework.¹²⁶ At the very least, it appears that principals need stronger training on how observations should be used for both evaluation and helping teachers improve their practices.¹²⁷ So, while teacher evaluation systems that incorporate VAM and observational components offer promise in using evidence to identify teacher performance, much more work needs to be done to make teacher evaluation and accountability systems identify teacher effectiveness as intended.

B. Teacher Motivation

In addition to being based on the notion that better teacher evaluation is critical for helping administrators to make better personnel decisions about teachers, teacher evaluation and accountability reforms are aimed at motivating teachers. As the logic of these reforms goes, teachers will perform better if they are properly motivated by evaluation and accompanying incentives. While this logic makes sense at first glance, there is little or no evidence to support it. Given the variety of factors that influence student progress, teacher evaluation and accountability reforms can discourage teachers from wanting to work in schools with poor and minority students.¹²⁸ As a result, these reforms could reinforce the common practice of assigning inexperienced teachers to these students and exacerbate incentives for schools to push out these students.¹²⁹

Teacher evaluation reforms also increase oversight over teachers and narrow the scope of teachers' autonomy, which in turn has the great potential to influence teachers' motivation negatively.¹³⁰ By increasing oversight through evaluation, these laws place more formal emphasis on teaching particular content and engaging in particular instructional practices

¹²⁴ *Id.* at 18.

¹²⁵ *Id.* at 14 (Figure 2).

¹²⁶ *Id.*

¹²⁷ See *id.* at 13.

¹²⁸ See BAKER ET AL., *supra* note 102, at 4.

¹²⁹ See Rachel O'Brien, Stanford Researchers Identify Troubling Patterns of Teacher Assignments Within Schools, STANFORD REP. Apr. 23, 2013, <http://news.stanford.edu/news/2013/april/school-teacher-assignment-042313.html>.

¹³⁰ See RICHARD M. INGERSOLL, WHO CONTROLS TEACHERS' WORK? 223 (2003).

toward specified objectives. To be sure, standards and assessments have specified for decades what students should know and be able to do (and thus have had strong implications for how and what teachers should teach). But direct evaluation of teacher performance through particular tests and observation rubrics constitutes a formal mechanism for specifying what teachers must do and narrows notions of what constitutes teaching to an even greater extent. Such policies further curtail teachers' autonomy to the extent that the evaluation is tied to formal accountability mechanisms, such as those involving tenure decisions, by intensifying the focus on particular objectives.

Research on job and work design indicates that self-determinism plays an important motivational role in work and especially teaching.¹³¹ Historically, teachers have placed high value on discretion, self-determinism, and authority over classroom work.¹³² Teachers consistently have identified restrictions to their decision-making and failure to provide adequate support for job performance as some of the greatest hindrances to their success.¹³³ Such restrictions include student testing demands, lack of adequate materials, resources, and facilities, and lack of support from administrators.¹³⁴ As such, increased levels of public oversight in education often have been marked by a focus on outputs and effectiveness that is more commonly associated with private sector management techniques for work that does not involve high skill levels or motivation that transcends the terms of employment.¹³⁵ External controls, such as high stakes testing and evaluation policies, accordingly tend to reduce teacher control and self-determinism and shift the focus of their work in ways that may reduce its perceived meaningfulness, the meaningful success of students, and the intrinsic rewards that teachers highly value.¹³⁶ Indeed, the amount of control teachers' perceive that they have over their work is related significantly to their engagement in work and their likely retention.¹³⁷ In short, although recent teacher evaluation and accountability reforms are partly aimed at increasing teachers' motivation, they have the strong potential to reduce it

¹³¹ See generally David B. Greenberger et al., Responses to Inadequate Personal Control in Organizations, 47 J. SOC. ISSUES 111 (1991) (discussing employee reactions to organizational structures).

¹³² See CUBAN, *supra* note 25, at 272–90.

¹³³ See INGERSOLL, *supra* note 130, at 223.

¹³⁴ See Cynthia Kopkowski, Why They Leave, NEA TODAY (April 2008), <http://www.nea.org/home/12630.htm>.

¹³⁵ RAYMOND E. CALLAHAN, EDUCATION AND THE CULT OF EFFICIENCY: A STUDY OF THE SOCIAL FORCES THAT HAVE SHAPED THE ADMINISTRATION OF THE PUBLIC SCHOOLS 178, 218 (2nd prtg. 1964).

¹³⁶ Linda Darling-Hammond & Elle Rustique-Forrester, The Consequences of Student Testing for Teaching and Teacher Quality, 104 Y.B. NAT'L SOC. STUDY EDUC. 289, 309–11 (2005), available at <http://eric.ed.gov/?id=EJ885537>.

¹³⁷ INGERSOLL, *supra* note 130, at 203–04, 223.

inadvertently, especially toward the students that need effective teachers the most.

C. Policy Coherence

Recently enacted teacher evaluation and accountability reforms raise significant concerns grouped around incoherence – the tensions, dilemmas, and conflicts that result when different federal, state, and district education reform policies and programs that are not aligned converge at the school and classroom levels. Especially since the U.S. Supreme Court decided *Brown v. Board of Education* in 1954, education policies have proliferated at the federal, state, and local levels, and the rate of proliferation has accelerated substantially in the past 20 years.¹³⁸ Schools now face a growing “barrage of demands” from several sources, including federal and state governments, local school boards, unions, and community groups.¹³⁹ These demands focus on several facets of schooling, ranging from curriculum and instruction, the use of time, testing, accountability, management, parent involvement, and many aspects of teacher workforce policy, including initial preparation, certification and licensure, evaluation, compensation, professional development, tenure, and retention. While demands from various policies originating from different levels of the education system may provide new opportunities for improvement, they also pose serious risk ultimately for improving and equalizing students’ learning opportunities.¹⁴⁰

For example, early research on the individual and aggregate effects of multiple federal categorical programs on school and district operations found that policy incoherence was associated with substantial program “interference” and “cross-subsidy,” which compromised the implementation and effects of these programs.¹⁴¹ These programs not only interfered with each other, they interfered with basic school-level operations and instructional programs. Moreover, schools and districts found that they had to reallocate resources from categorical programs and from general operating funds to meet implementation demands. More recently, research has found that “cluttered and contradictory” federal, state, and district

¹³⁸ SUPERFINE, *supra* note 9, at 201–02.

¹³⁹ Meredith L. Honig & Thomas C. Hatch, *Crafting Coherence: How Schools Strategically Manage Multiple, External Demands*, 33 *EDUC. RESEARCHER* 16, 16 (2004).

¹⁴⁰ See David K. Cohen & James P. Spillane, *Policy and Practice: The Relations Between Governance and Instruction*, in *DESIGNING COHERENT EDUCATION POLICY: IMPROVING THE SYSTEM* 35, 76–78 (Susan H. Fuhrman ed., 1993).

¹⁴¹ U.S. DEP’T OF EDUC., *THE AGGREGATE EFFECTS OF FEDERAL EDUCATION PROGRAMS* 5–6 (1981), available at <http://www.rand.org/content/dam/rand/pubs/reports/2005/R2638.pdf>.

policy initiatives are associated with fragmentation and compromise improvement at the school level.¹⁴²

On a very broad level, some aspects of recent teacher evaluation laws are likely to promote coherence. Many intersecting state and federal laws that focus on teacher evaluation were developed in policy context focused on the improvement and evaluation of teachers, particularly in ways that are linked to the Common Core State Standards. Moreover, a few states (like Louisiana and Texas) have begun to align evaluations for teacher preparation programs with evaluations for practice teachers.

However, significant concerns grouped around incoherence remain for teacher evaluation and accountability reforms. On a fundamental level, conflicts and tensions in the policy landscape arise because at least some of these different policies are grounded in conflicting conceptions of what improves student learning in schools. For example, a recently enacted law in Illinois restructuring requirements for principal certification was driven by a view of school leadership that identifies instructional quality as a property of schools as organizations rather than a property of individual teacher talent.¹⁴³ This vision foregrounds the capacity of school leaders to lead effective professional learning communities marked by high levels of trust, work productively with parents and community agencies, and meet the social and emotional learning needs of students and adults in their schools. In contrast, the standards-based, assessment-driven vision of Illinois' teacher evaluation and accountability system conceptualizes the solution to underperforming schools as one of improving teacher talent by motivating individual teachers through external accountability measures.¹⁴⁴

Teacher evaluation and accountability reforms also raise problems of incoherence because of issues to which they do not attend. Other than provisions that directly align particular state, local, and school functions to state standards and test scores, these laws are largely silent on the quality of particular workforce development functions. These functions include work done by teachers and principals; the development of teachers' knowledge, especially around student performance evidence; how multiple workforces, e.g. the teacher, principal, and district-level administrator workforces, should be related; and the workplace conditions of schools and school districts that affect teacher performance. Similarly, these laws have little to

¹⁴² Fred M. Newmann et al., *Instructional Program Coherence: What It Is and Why It Should Guide School Improvement Policy*, 23 *EDUC. EVALUATION & POL'Y ANALYSIS* 297, 299 (2001), available at <http://epa.sagepub.com/content/23/4/297.full.pdf>.

¹⁴³ 105 ILL. COMP. STAT. ANN. 5/21-7.6 (West 2012) (repealed by its own terms 2013).

¹⁴⁴ 105 ILL. COMP. STAT. ANN. 5/2-3.25g (West Supp. 2012), amended by 2013 Ill. Legis. Serv. 513 (West).

do with curriculum, or professional development more broadly, or other education policy efforts that may be underway in a state, such as the expansion of charter schools or alternative routes to teacher certification. In short, teacher evaluation and accountability reforms have not been sufficiently designed to converge around goals, strategies, and resource demands underlying many other federal and state laws. Given their increasing importance in education policy, these reforms accordingly have the strong potential to exacerbate pre-existing policy incoherence and ultimately weaken students' learning opportunities.

D. The Concept of Teaching

On their most fundamental level, many recently enacted teacher evaluation and accountability reforms reflect an underlying concept of teaching that may ultimately impair students' learning opportunities. Given the centrality of VAM in teacher evaluation and accountability laws, they reflect an understanding of effective teaching as one that can be measured well through standardized tests. Current tests used for VAM, however, are quite narrow insofar as they focus on basic skills and primarily employ multiple-choice questions.¹⁴⁵ When teachers teach to the content embedded in such tests, they may underemphasize or ignore important activities such as writing, inquiry, and complex problem solving.¹⁴⁶ Indeed, teachers' measured effectiveness can differ significantly depending on which test is used.¹⁴⁷ In particular, teachers who are rated highly through VAM and focus on basic skills generally fail to be rated highly using conceptually oriented tests.¹⁴⁸

Such findings point to the broader problems of enhancing bureaucratic controls to improve teaching and learning. As discussed above, the focus on testing increases oversight over teachers and narrows the scope of teachers' autonomy. However, teacher autonomy is not only beneficial for teacher motivation; it also is a central factor influencing teachers' abilities to provide strong educations to students. The nature, goals, and "technologies" of teachers' work—teaching and learning—are ambiguous, often unpredictable, sometimes contested, and highly contextual.¹⁴⁹ The work is

¹⁴⁵ DARLING-HAMMOND, *supra* note 118, at 20.

¹⁴⁶ *Id.*

¹⁴⁷ JESSE ROTHSTEIN, NAT'L EDUC. POL'Y CTR., REVIEW OF LEARNING ABOUT TEACHING 5 (Kevin Welner ed., 2011) (reviewing MEASURES OF EFFECTIVE TEACHING PROJECT, LEARNING ABOUT TEACHING), available at <http://nepc.colorado.edu/files/TTR-MET-Rothstein.pdf>.

¹⁴⁸ DARLING-HAMMOND, *supra* note 118, at 20.

¹⁴⁹ See generally, CUBAN, *supra* note 25.

dynamic, idiosyncratic, and non-routine.¹⁵⁰ Teachers are expected to achieve multiple objectives at the same time, including serving as facilitators of student learning, role models for students, and partners with parents and communities.¹⁵¹ As Dan Lortie argued in his seminal book *Schoolteacher*, teaching is fundamentally “people work,” and teachers must “adjust and readjust their actions in line with hoped-for outcomes; they must monitor their steps and make corrections as they proceed.”¹⁵² Public statements emerging from the 2012 teachers’ strike in Chicago reflects such an underlying conception of teaching. As the CTU argued, Illinois’ teacher evaluation and accountability reforms will deskill teachers and harm students.¹⁵³ As student populations become more diverse and demands for student learning evolve, greater autonomy may be even more necessary.

To be sure, teachers should not be accorded unlimited autonomy. Complete autonomy is neither practically feasible nor likely to result in consistently improved instructional practices. A certain degree and type of autonomy, however, is critical for allowing teachers to tailor instruction to the local conditions of their schools and classrooms. With such “bounded autonomy,” teachers are likely to focus less on teaching to the test and more on providing high quality, comprehensive instruction.¹⁵⁴

IV. MOVING FORWARD WITH TEACHER EVALUATION AND ACCOUNTABILITY

Teacher evaluation and accountability reforms involve significant promise and pitfalls. Teachers are one of the most important factors inside schools that influence students’ learning opportunities and achievement. They are especially important for structuring the learning opportunities of poor and minority students. Yet, teacher evaluation and accountability traditionally has played a minimal role in ensuring that all students are exposed to high quality and effective teachers. Indeed, despite the quality of their actual performance and effectiveness in providing all students with

¹⁵⁰ See Brian Rowan et al., *Teaching as a Nonroutine Task: Implications for the Management of Schools*, 29 *EDUC. ADMIN. Q.* 479, 481–82 (1993).

¹⁵¹ Mark A. Smylie et al., *The Work of Teachers*, in 1 *21ST CENTURY EDUCATION: A REFERENCE HANDBOOK 4-6* (Thomas L. Good ed., 2008).

¹⁵² DAN C. LORTIE, *SCHOOLTEACHER: A SOCIOLOGICAL STUDY* 135, 137 (with a new preface 2002).

¹⁵³ Cf. Press Release: CPS Fails to Negotiate Fair Contract to Prevent First Strike in 25 Years, CHICAGO TEACHERS UNION (Sept. 9, 2012), <http://www.ctunet.com/blog/cps-fails-to-negotiate-fair-contract-to-prevent-first-labor-strike-in-25-years>.

¹⁵⁴ Cf. Catherine S. Durso, *An Analysis of the Use and Validity of Test-Based Teacher Evaluations Reported by the Los Angeles Times: 2011*, NAT’L EDUC. POL’Y CTR. ii (2012), <http://nepc.colorado.edu/files/rb-latimesii.pdf> (explaining that standardized tests, if used as the sole criterion for gauging a teacher’s performance, will lead to “teaching to the test” rather than “high-quality, comprehensive instruction.”).

strong learning opportunities, teachers are usually evaluated as performing satisfactorily. Recent teacher evaluation reforms are aimed at remedying the problem of evaluation inflation. By evaluating teachers largely on the basis of observations and increases in the standardized test scores of their students, and holding teacher accountable for their performance, these reforms hold the potential of impacting one of the most important factors in students' educations. These recently enacted reforms, however, raise a host of interlocking problems. They are riddled with psychometric issues that weaken their validity, may fail to increase and may possibly decrease teacher motivation, raise issues of policy incoherence, and reflect an underlying concept of teaching that may weaken overall teacher effectiveness.

Given such problems, how should laws focused on teacher evaluation be modified moving forward? This section offers three recommendations for helping us capitalize on the opportunities that teacher evaluation and accountability reforms offer: (1) develop a robust conception of an effective teacher, (2) rebuild teacher evaluation and accountability systems with this conception as their foundation, and (3) align teacher evaluation and accountability reforms with other education policies.

A. Recommendation One: Develop a Robust Conception of an Effective Teacher

First, and most importantly, teacher evaluation and accountability reforms should be grounded in a strong conception of what constitutes an effective teacher. This conception should be guided by the best available evidence and thus should focus on what it takes to provide all students with high quality learning opportunities. On a very general level, this conception should not simply include helping students perform well on narrowly constructed, multiple choice assessments. It should also involve providing students with more conceptually oriented learning opportunities. Notably, the concept of an effective teacher would differ by discipline – while “effective” teaching in mathematics and language arts share many broad characteristics (such as student engagement), the content of the disciplines requires the vision of effective teaching to be different in important ways.

A brief example of the type of knowledge an effective mathematics teacher could have illustrates how a discipline-specific concept of such a teacher could be developed. Mathematics education researchers increasingly have emphasized the need for teachers to develop “mathematical knowledge for teaching”. This type of knowledge actually includes two different types of knowledge: (1) pedagogical content

knowledge, and (2) subject matter knowledge.¹⁵⁵ Drawing from early research on teacher knowledge,¹⁵⁶ pedagogical content knowledge includes knowledge of the interrelated aspects of content and students (e.g. deciding which of several errors students are likely to make based on where they are in the development of their mathematical understanding); content and teaching (e.g. knowing instructionally viable models for place value and how to teach them effectively); and content and curriculum (e.g. understanding where division with whole numbers fits within the elementary school curriculum in relation to students' developmental trajectories).¹⁵⁷ Taken as a whole, pedagogical content knowledge is, at first, an understanding of the prior conceptions and knowledge that students of different ages and backgrounds bring with them into the classroom. To apply this knowledge successfully, a teacher must communicate mathematics in ways that students can learn effectively.

Subject matter knowledge, on the other hand, includes two separate domains of knowledge: common content knowledge and specialized content knowledge.¹⁵⁸ Common content knowledge is the substantive type of mathematical knowledge that is used by both teachers and those in other professions. For example, common content knowledge includes understanding how to use percentages to compute amounts of discounts. Specialized content knowledge is the type of substantive knowledge required exclusively for teaching. Such knowledge includes an understanding of how to evaluate the validity of the mathematics in solution methods.¹⁵⁹ Indeed, researchers have found that both common and specialized content knowledge are related to increased student achievement.¹⁶⁰ As such, it is not enough for teachers to just know mathematics. Rather, teachers should know mathematics in the ways that are needed specifically for teaching students. Strong mathematical knowledge for teaching, however, is clearly not sufficient to make an effective teacher. Teachers should also know how to generate student interest, engage with parents and communities, and productively work with

¹⁵⁵ Deborah Loewenberg Ball et al., Content Knowledge for Teaching: What Makes It Special?, 59 J. TCHR. EDUC. 389, 402–03 (2008).

¹⁵⁶ See Lee S. Shulman, Those Who Understand: Knowledge Growth in Teaching, 15 EDUC. RES. 4, 9 (1986).

¹⁵⁷ Deborah Loewenberg Ball et al., Content Knowledge for Teaching: What Makes It Special 4–5, available at <http://conferences.illinoisstate.edu/nsa/papers/thamesphelps.pdf> (last visited Oct. 19, 2013).

¹⁵⁸ Id.

¹⁵⁹ See generally DEBORAH LOWENBERG BALL & HYMAN BASS, Interweaving Content and Pedagogy in Teaching and Learning to Teach: Knowing and Using Mathematics, in MULTIPLE PERSPECTIVES ON THE TEACHING AND LEARNING OF MATHEMATICS 83–104 (Jo Boaler ed., 2000).

¹⁶⁰ See, e.g., Heather C. Hill et al., *Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement*, 42 AM. EDUC. RESEARCH J. 371, 371–406 (2005) (discussing the correlation between teacher knowledge in a specific subject and student achievement).

other teachers and principals. However, mathematical knowledge for teaching potentially constitutes a key element for providing students with robust learning opportunities in mathematics. It is precisely this sort of effective teaching that requires a formation of the conceptual foundation of teacher evaluations and accountability reforms.

Developing this sort of conception is difficult. Current political forces push policymakers toward quick fixes in education with demonstrable results, such as fast and clear increases in student standardized test scores. While such indicators are important, they reflect only a narrow conception of effective teaching. In order to articulate such a conception of effective teaching in various disciplines, states should convene representatives from major educational stakeholders, such as teachers, school administrators, policymakers, and researchers. While some states have made efforts to reach out to such stakeholders, these efforts should be continued and bolstered as research-based visions of effective teaching are more explicitly articulated. In doing so, policymakers can lay the groundwork for much stronger teacher evaluation and accountability systems.

B. Recommendation Two: Rebuild Teacher Evaluation and Accountability Systems Using a Robust Conception of an Effective Teacher

As discussed throughout, teacher evaluation and accountability systems offer much promise but need to be rebuilt if they are to result in increases in learning opportunities for all students. Most importantly, these systems need to be grounded in a strong conception of an effective teacher. Once such framework is developed, the nuts and bolts of the evaluation system should align with this concept. The idea of mathematics knowledge for teaching, as discussed under Recommendation One, can again serve as an instructive example for the foundation of teacher evaluation and accountability systems.

Given such a foundation, teacher evaluation systems will need to be restructured dramatically. Current systems that rely on VAM are too narrow to account for this kind of robust vision. Further, many of the observation rubrics used to implement teacher evaluation systems are not discipline specific. Moreover, traditional assessments of teachers' mathematical knowledge are inappropriate for assessing mathematical knowledge for teaching because they focus on subject matter knowledge or pedagogical knowledge as separate constructs. As part of the Learning Mathematics for Teaching ("LMT") Project, however, researchers at the University of Michigan have developed a suite of assessments that measures elementary

school teachers' mathematical knowledge for teaching in different areas of mathematics.¹⁶¹ These areas include numbers and operations, algebra, and patterns and functions. Items on these assessments not only capture whether teachers can answer the problems that they use with students, but also how teachers solve the special mathematical tasks that arise during teaching. For example, items on the LMT assessment include the following questions: (1) Is the number 0 even or odd? (2) Can two different triangles have the same area? (3) Given three different multiplication strategies, which method can be used to multiply any two numbers? Although these assessments are not currently used in teacher evaluation systems, they are increasingly being used in various professional development settings around the United States. These assessments represent a substantial advance over current teacher knowledge assessments and indicators, such as "basic skills" tests given to teacher candidates in some states and whether teachers have a degree in their field.

To be sure, such assessments focus only on what teachers know and not how well they perform in the classroom. Because what teachers actually do with this knowledge is crucial as well, evaluations of teachers' classroom performance should be part of such an assessment. Discipline specific observation rubrics that focus on the extent to which teachers engage in instructional practices in line with the underlying conception of an effective teacher should be developed and implemented. Evaluators should be trained to use these rubrics to evaluate teachers effectively and consistently, and ultimately provide strong feedback to teachers to help them improve. If any form of VAM is to be used, test developers should provide strong evidence that the assessment practices are valid and, more specifically, that the assessments evaluate teacher effectiveness in line with a robust conception of effective teaching and not a range of other factors.

On a more fundamental level, any student assessments used as part of teacher evaluation should focus on a vision of learning that is much more robust than what is currently reflected under traditional standardized assessments. For example, the Gordon Commission on the Future of Assessment in Education, established by Educational Testing Service and composed of researchers, policymakers, and practitioners, made several observations and recommendations about student assessments.¹⁶² Chief among these observations is that at their core, assessments serve as

¹⁶¹ See generally Heather C. Hill et al., *Developing Measures of Teachers' Mathematics Knowledge for Teaching*, 105 *ELEMENTARY SCHOOL J.* 1, 11–30 (2004).

¹⁶² A Public Policy Statement, THE GORDON COMMISSION ON THE FUTURE OF ASSESSMENT IN EDUCATION, www.gordoncommission.org/rsc/pdfs/gordon_commission_on_technical_report.pdf (last visited Oct. 19, 2013).

statements about what educators, policymakers, and parents want their students to learn. Current assessments generally fail to assess whether students can evaluate the validity and relevance of different pieces of information and draw conclusions from them, make conjectures and seek evidence to test them, contribute to their job or community networks, and generally make sense of the world. The Gordon Commission has recommended that assessments should not only serve as instruments for accountability, but that the assessments should also serve as “tools that provide teachers with actionable information about their students and their practice in real time” and be “designed an used to support high-quality education.”¹⁶³

This recommendation surrounding the purpose of student assessments beyond accountability is critical. In order for such assessments to motivate teachers to engage in different instructional practices, teachers must perceive them as well aligned with high quality instructional practices. Moreover, teacher evaluation and accountability systems must be perceived as a tool for improving instruction and developing teachers’ capacities instead of simply incentive mechanisms. Along similar lines, teacher evaluation and accountability systems should be designed with the importance of local context in mind. Because the nature, goals, and technologies of teaching are ambiguous, unpredictable, and highly contextual, evaluation systems (and evaluators) should be flexible enough to account for a range of teacher behavior. If teachers are to be granted autonomy to teach effectively, evaluation and accountability systems should account for this autonomy as well. This is not to say that these systems should allow for any type of behavior to be rated as satisfactory. Such minimal evaluation and accountability processes drove the push for teacher evaluation reform in the first place. Instead, these systems should be grounded in the notion that teachers should be given a certain degree and type of autonomy, with its boundaries set by a strong conception of an effective teacher. With such an approach, especially if focused not just on incentivizing teachers but helping them improve, teacher evaluation and accountability systems would be much more likely to improve learning opportunities for all students.

¹⁶³ Id. at 5–6.

C. Recommendation Three: Align Teacher Evaluation and Accountability Reforms with Other Policies

The incoherence in states' approaches to education policy is rarely raised in political discussions. It is, however, another fundamental issue that ultimately weakens students' educational opportunities. If states are to improve the coherence of their education policies, they must not focus exclusively on improving particular policy areas, such as teacher evaluations and accountability, but on improving coherence across education policy areas as well. In order to improve coherence, states should commit to a vision of teaching and learning as an "anchor" for other major policies. Such policies should include those governing the spectrum of teacher workforce functions. These functions include initial preparation, recruiting, development, remediation, compensation, retention, and firing. Moreover, such policies should include those that are not directly part of teacher workforce management, such as curriculum and student assessment.

Several elements should be considered for constructing this anchor vision of teaching and learning. The conception of an effective teacher developed under Recommendation One should serve as a key element of this vision. States also should attend to other major visions of teaching and learning that are already present and potentially guiding education policies. As discussed above, every state currently has a set of student learning standards in place, which theoretically serve as a basis for a range of other policies, such as those governing curriculum and testing. Moreover, 45 states and the District of Columbia use phasing consistent with the Common Core of State Standards. Promising theories of effective schools and school improvement, such as those focused on the development of practicing teachers' and principals' knowledge and skills, could constitute other key elements. Indeed, the conception of an effective teacher constructed under Recommendation One should be aligned with these subsequent elements. As it stands, state standards have created some areas of overlap across different reforms. Many gaps remain, however, especially as new reforms are constantly enacted. Laws governing education reform should work together and be mutually reinforcing, no matter what underlying vision is in place.

Based on such a vision, states should modify their laws to align with this vision. In doing so, states should focus on developing the capacities, processes, leadership, and incentives for schools to craft and manage coherence.¹⁶⁴ This structure should focus on putting schools in strong

¹⁶⁴ See generally Honig & Hatch, *supra* note 139.

positions to balance external legal demands with local goals and strategies. Moreover, state laws should frame the role of the state department of education as an institution that develops and manages coherence at the local level. State departments of education currently spend a substantial amount of time and resources in assisting localities in searching for and using information. Instead, state departments should promote the vision of learning and teaching developed by the state and built into law as an anchor around which local practices align. The state departments should also hold schools accountable for aligning their programs and practices with these anchors. Further, the departments should negotiate with legislative and regulatory bodies when contradictory demands are made. By making states' governance role more robust but focused on central goals rather than local processes, coherence would be improved without significantly weakening local discretion.

V. CONCLUSION

The wave of laws introducing teacher evaluation and accountability reforms represents the most recent major effort to improve students' learning opportunities. Education reformers long have focused on traditional educational inputs to improve these opportunities, but their efforts have yielded inconsistent success at best. Teacher evaluation and accountability reforms have the potential to start implementing the promises of these earlier efforts. These reforms focus on teachers as a critical factor structuring students' learning opportunities, and they directly engage with what teachers do and what students learn. However, as these reforms are currently structured, they are riddled with problems. Reforming teacher evaluation and accountability reform is a challenging process. It requires deep engagement from a range of stakeholders about what an effective teacher is, and this discussion is ultimately rooted in what and how students should learn. It also requires a commitment to looking deeply at evidence about assessment, accountability, teaching, and learning. Lastly, it requires the dedication to look not only at one type of education reform but all of them. Such conversations are difficult to have productively in the current political climate. Such conversations are necessary, however, if all students are to be provided with the educational opportunities that they deserve.