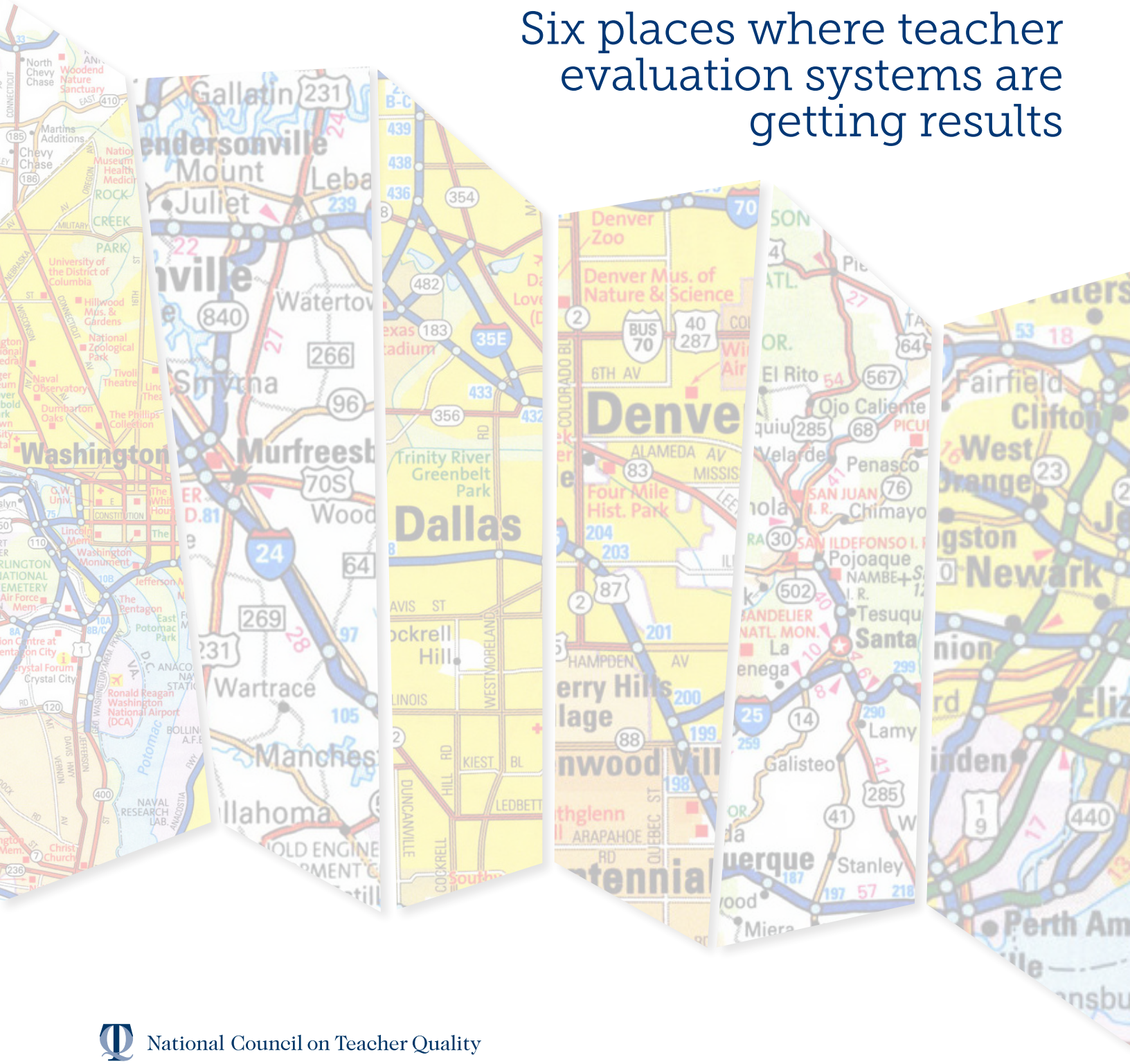
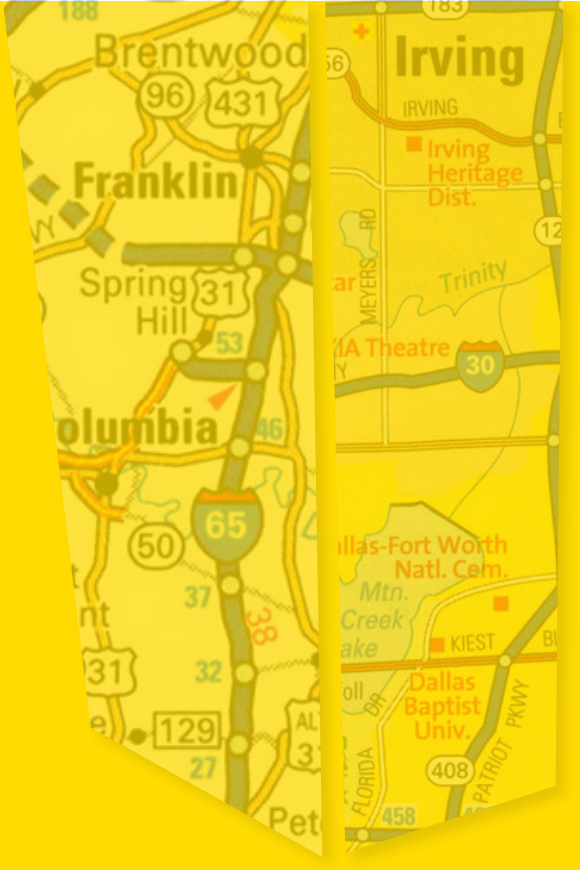
Making a difference

Six places where teacher evaluation systems are getting results





Authors: Hannah Putman[1], Elizabeth Ross, and Kate Walsh

Analysts: Kelli Lakis and Kency Nittler

NCTQ receives all of its funding from foundations and private donors. We thank them, and specifically, The Joyce Foundation and the Daniels Fund, for the generous and sustained support of our work.

We also thank the following individuals for reviewing relevant portions of this analysis. Their insight and expertise have been valuable and we are grateful for both.

Betsy Press, *Former Deputy Chief of Impact, District of Columbia Public Schools*

Christopher Eide, *Director, Educator Quality, New Mexico Public Education Department*

Christopher Lewis, *Deputy Chief, IMPACT, District of Columbia Public Schools*

Dorothy Smith, *Managing Director, Best in Class, The Commit Partnership*

John Vega, *Deputy Chief, Human Capital Management, Dallas Independent School District*

Larisa Shambaugh, *Former Chief Talent Officer, Newark Public Schools*

Matt Montaño, *Former Deputy Cabinet Secretary, Teaching and Learning, New Mexico Public Education Department*

Nicole Wolden, *Director of Growth and Performance, Denver Public Schools*

Paul Fleming, *Assistant Commissioner of Teachers and Leaders, Tennessee Department of Education*

Sarah Almy, *Executive Director, Talent Management, Denver Public Schools*

Suzy Smith, *Manager of Performance Management, Dallas Independent School District*

Todd Williams, *Chairman and CEO, The Commit Partnership*

Overview

Teacher evaluation has undergone rapid reform over the past decade.[2] School systems set out to help strengthen and improve their existing evaluation systems, which were just not functioning effectively. Under past systems, almost all teachers were being rated as satisfactory and therefore largely viewed as interchangable.[3]

Evaluation systems must be able to differentiate among different teachers' effectiveness and apply that information in key personnel decisions in order to build the strongest possible teacher workforce. An examination of many states' and school districts' current approaches to evaluation shows that they have made some real progress in developing improved systems over the past decade. For example, many teacher evaluation systems historically included only two rating levels — satisfactory and unsatisfactory — with almost all teachers earning the former.[4] Today, most states' and large districts' evaluation systems consist of three or more rating levels. While previous ratings were based primarily on observations of the teacher by a principal, more evaluation systems now factor in objective measures of student growth, as well as other measures such as student input in the form of surveys. Whereas many school systems once routinely evaluated tenured teachers on a two- to five-year cycle, approximately half of the 100 largest districts and half of all states now require annual, summative evaluations for all teachers and have policies to use the results of evaluation systems to inform teacher compensation.[5,6]

Despite this clear progress, strengthening evaluation systems so that they reflect the genuine distribution of teacher talent in a district or state has proven challenging. One of the setbacks was the implementation of new teacher evaluation measures in tandem with new college- and career-ready standards and assessments for students. These simultaneous transitions meant that teachers were making significant changes in what they taught at the same time they were being more rigorously evaluated, causing some teachers to bristle at these policy changes.

Fortunately, we can identify six pioneers — four districts and two states — which are reporting genuine successes. They are implementing many of the same components commonly found in many state and district systems; however, their results are setting them apart.

These six systems are distinguishing among teachers of varying quality and are delivering the changes initially sought by districts and states across the nation. Some of these systems are already able to boast measurable evidence that the teacher workforce is improving, with higher-performing teachers staying longer and weaker teachers, who previously might never have even known they were weak, choosing to leave.

This analysis focuses on these six systems: Dallas Independent School District, District of Columbia Public Schools, Denver Public Schools, Newark Public Schools, New Mexico, and Tennessee. It depicts how evaluation systems can benefit teachers and, most importantly, students.

The districts and states discussed can each report publicly available evidence of the positive impacts of their evaluation systems. Although the specific successes observable in the districts and states profiled here differ, each offers examples of progress that is possible under well-designed and well-implemented teacher evaluation systems which achieve a more honest distribution of teacher talent and which emphasize continuous improvement. The case studies are drawn from official policies, reviews of the evaluation system that these states and districts have conducted, independent studies, and interviews with district and state staff.

Goals of teacher evaluation

Strong teacher evaluation systems, when paired with supports and incentives, are designed to do the following:

1) Provide a more valid measure of teacher quality by distinguishing between teachers at different performance levels;
2) Recognize strong teachers and keep them in the classroom;
3) Encourage consistently less effective teachers to leave the classroom;
4) Help all teachers improve;
5) Recruit more effective new teachers; and
6) Achieve gains in student learning and other positive student outcomes.

What these six systems have in common

The knowledge base for building a strong evaluation system is still young and is continuously being refined, but some clear principles of strong practice have emerged from a decade of innovation and implementation (see the appendix for more details on these practices and supporting research). As is immediately evident, the districts and states highlighted here have each implemented many of these strong practices.

Their success is a result of adherence to core principles[7]

Strong Practice	Dallas Independent School District	Denver Public Schools	District of Columbia Public Schools	Newark Public Schools	Tennessee	New Mexico
Multiple measures	✓	✓	✓	✓	✓	✓
Student surveys	✓	✓	✓		✓	✓
Objective measures of student growth	✓	✓	✓	✓	✓	✓
At least three rating categories	✓	✓	✓	✓[8]	✓	✓
Annual evaluations and observations for all teachers	✓	✓	✓	✓	✓	✓
Professional development tied to evaluation	✓	✓	✓	✓	✓	✓
Written feedback after each observation[9]	✓	✓[10]		✓	✓	✓

Although 89 of the 100 largest districts use multiple measures, only 21 use or allow student surveys.[19]

One notable feature common among the six locales highlighted here is the use of multiple measures to comprise the overall evaluation rating. The systems vary in the precise measures used, but each uses at least three measures and includes some measures of student learning, as well as observations and, in many cases, student surveys.

Their success is built on a thoughtful approach to weighting individual evaluation components

Weight of individual components (vary by grade and subject)	Dallas Independent School District	Denver Public Schools	District of Columbia Public Schools	Newark Public Schools[11]	Tennessee	New Mexico
Observations[12]	50-80%	30-35%	30-75%	55-85%[13]	50% (all qualitative measures, including student surveys)	40%
Student achievement or growth	20-35%	50%	15-50%	15-45%	50%	35%
Student surveys	0-15%	0-10%	0-10%	0%	See "Observations"	5%
Professionalism	0%	10-15%	0%	0%	0%	15%
Commitment to the school community	0%	0%	10%	0%	0%	0%
Teacher attendance	0%	0%	0%	0%	0%	5%

Evaluation systems can be an essential part of a district's or state's talent management strategy, not only to give teachers feedback and support, but also to inform myriad decisions such as eligibility for leadership roles, raises, or retention in the classroom.

Their success is made possible by linking evaluation to key personnel decisions

Key decisions	Dallas Independent School District	Denver Public Schools	District of Columbia Public Schools	Newark Public Schools	Tennessee	New Mexico
Ties compensation to evaluations	✓	✓	✓	✓	✓	✓
Selects cooperating teachers to host student teachers based on evaluations	✓	✓			✓	
Selects teachers for leadership opportunities based on evaluations	✓	✓	✓	✓		✓
Tailors professional development based on evaluations	✓	✓	✓[14]	✓	✓	✓
Makes teacher dismissal decisions based on evaluations	✓	✓	✓	✓	✓	
Incentivizes effective teachers to work in high-need schools	✓	✓	✓	✓		
Other	Identifies effective teachers to teach summer school	Earning or losing non-probationary (or tenure) status[15]	Adapts evaluation rubric to assess teacher applicants during hiring process		Informs teacher preparation program accountability; influences licensure advancement and renewal decisions; determines if a new teacher qualifies for tenure; informs layoff decisions	Informs teacher preparation program accountability; influences licensure advancement and renewal decisions; informs selection of the state Teacher of the Year

Conversations with district and state staff in these six sites emphasized that attaching meaningful consequences to evaluation systems encourages teachers and principals to take them seriously. Teachers were more inclined to seek support, and principals were more motivated to have difficult conversations with their staff. While teachers continue to earn bonuses or raises for many reasons, compensation tied to positive ratings on evaluations is increasingly common.

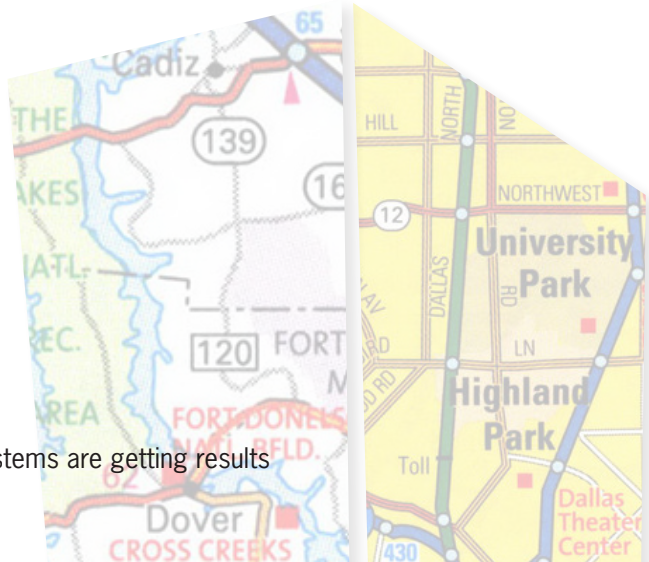
Their success is fueled by making strategic compensation decisions

	Dallas Independent School District	Denver Public Schools	District of Columbia Public Schools	Newark Public Schools	Tennessee	New Mexico
Higher evaluation rating	Raise	Raise	Raise & bonus	Raise & bonus	Raise	Award
Teaching a hard-to-staff subject	Bonus	Bonus[16]	Bonus	–	Raise	Award
Teaching in a high-need school	Bonus	Bonus[17]	Bonus	–	Raise	Award
Cost-of-living adjustment	–	Raise	–	Raise	–	–
Additional year of teaching	–[18]	Raise	Raise	Raise	Raise	–
Earning an advanced degree	–	Raise	Raise	–	Raise	–
Other		Additional professional development; Leadership positions		Teacher leadership positions	Additional roles or responsibilities	

With the exception of Denver Public Schools, whose superintendent has remained in place since its evaluation system was first piloted, every system featured here has survived changes in leadership. While these systems continue to evolve and improve, they have all maintained their core principles.

The evaluation systems featured here include many of the characteristics that research and common sense suggest will yield the greatest benefits for teachers, school systems, and students. However, the staff implementing these systems emphasized that they have only reached this point because of a consistent commitment to assessing and improving their systems. None of these evaluation systems got everything right in the first year of implementation. System leaders gathered feedback from teachers, principals, and other stakeholders; analyzed the data they had gathered; and used this information to identify weak points and refine their systems. This systemic commitment to continuous improvement persevered despite leadership transitions experienced by each of the districts and states profiled here.

Evaluation is not a silver bullet that will automatically improve teacher effectiveness and student outcomes. But as the following six examples illustrate, with careful implementation and a commitment to build upon what works and remedy what does not, an evaluation system can be an essential tool in state and district efforts to ensure that every student has access to effective teachers.



Dallas Independent School District (DISD)

Evaluation System: Teacher Excellence Initiative (TEI)

History: Implemented in 2014 under Superintendent Mike Miles and continues under Superintendent Michael Hinojosa.

How it works: Teachers earn one of seven evaluation ratings that are then combined with other criteria (e.g., previous year's ratings, additional evaluation processes) which determines teachers' pay levels.

Impact on salary: Maximum: *$100,000 ($90,000 in salary, plus up to $10,000 in bonuses)*[20] *, achievable in as few as six years.*

Maximum prior to evaluation system implementation: *$87,338, achievable in 37 years with a Ph.D.*

REFLECTING ON SYSTEM

Notable features

- The district sets a target distribution for the percentage of teachers earning each rating. This policy forestalls the upward creep in ratings that may not be correlated with achievement and allows the district to budget responsibly.
- While all teachers are evaluated on core aspects of their job (classroom performance, students' achievement, students' classroom experiences), teachers must go through an additional review to reach the higher rating levels. Only teachers who are in the top 30 percent of their peer group[21] , are in at least their third year of service, and meet other performance criteria can apply to be in these elite categories.[22]

> Dallas is "able to identify our most effective teachers… [and we see that] retention of our most effective teachers is a strength of the system."
>
> — Suzy Smith, *Manager of Performance Management*[37]

- All teachers are formally observed several times a year, but the district also requires short "spot observations," which can save principals time and give them more real-time opportunities to engage with teachers and learn about their strengths and areas for growth.[23]
- To provide more salary stability and protect teachers from an anomalous bad year, Dallas averages a teacher's current and previous year's ratings to determine salary. Also, teachers' salaries only drop if they have earned a lower effectiveness level for three consecutive years.
- Teachers who earn one of the lowest two rating categories do not receive a raise.
- The Dallas evaluation system has a targeted distribution for all of its evaluation ratings, meaning that the district limits the number of teachers who fall into one of the ratings categories it has created. No more than 3 percent of teachers can receive the lowest rating, and no more than 2 percent can receive the highest rating.

Opportunities for improvement

- The system involves two different rating systems: one to measure performance in the most recent school year and a more cumulative system that incorporates annual ratings and other data to determine pay. These two systems have similar names, which has caused confusion.
- In a 2017 survey, roughly a quarter of teachers felt the evaluation system was unfair to teachers facing greater challenges in low-performing schools.[24]

Notable changes to system since inception

- When this system first started, salaries were based entirely on performance. Now, Dallas offers retention bonuses (which grow as teachers earn higher ratings) so that teachers earning all but the lowest ratings can see their compensation increase each year.
- The district plans to reduce the number of possible compensation levels, as well as increase the current cap on the percentage of teachers who can qualify for the highest rating.[25]

BUILDING BUY-IN

Steps to build buy-in

The district gathers feedback on the evaluation rubric from teachers, principals, and other stakeholders and calibrates it through campus walks and data analysis.[26] Dallas also recruits teachers on each campus to attend regular evaluation system trainings and to share this information with their colleagues.

Surveys of teachers and administrators

In a 2017 survey, only 41 percent of teachers were satisfied with the system, but this is an improvement from 2014-2015 when only 31 percent were satisfied.[27] Despite this level of dissatisfaction, many teachers like the individual pieces of the system. About three quarters of teachers found each aspect of the evaluation (spot observations, extended observations, and summative evaluation feedback) helpful.[28]

A 2017 survey found that administrators generally liked the system (72 percent were satisfied).[29]

EVIDENCE OF IMPACT

Recognizes strong teachers and keeps them in the classroom?

Yes, 98 percent of the highest rated teachers stayed in the district after the 2016-2017 school year.[30]

Encourages consistently less effective teachers to leave the classroom?

Yes, 50 percent of consistently unsatisfactory teachers left the district after the 2016-2017 school year.[31]

Helps all teachers improve?

Yes, teachers' average spot observation and overall evaluation scores are rising, suggesting that teachers may be performing better over time (although score inflation could also cause these increases).[32]

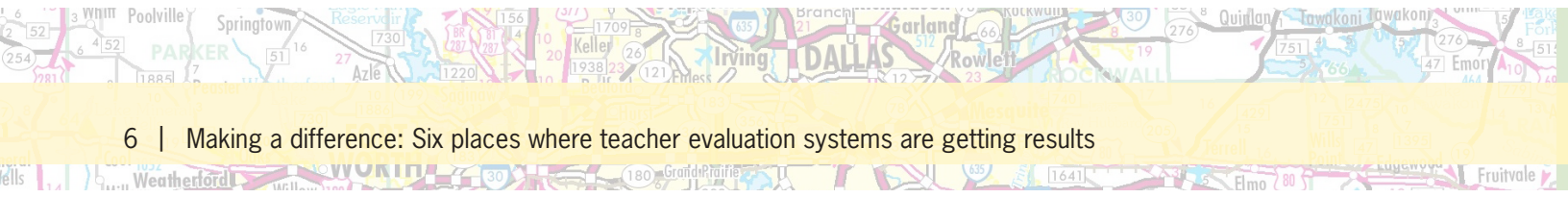
> Because of ongoing calibration training for evaluators, "we know that proficient is proficient, regardless of campus, content area, or grade level."
>
> — Suzy Smith, *Manager of Performance Management*[38]

Correlates with increased student achievement?

Yes, student proficiency across all subjects and grades (already on a modest upswing) increased more steeply by 7 percentage points from 2015 to 2017, closing the proficiency gap between Dallas and the state of Texas by 3 points.[33] Note that these gains cannot be tied directly to the evaluation system without a controlled study.

Other evidence of impact?

Dallas focused resources on a set of schools identified as "Improvement Required" by the state education agency. These resources focus heavily on strategic staffing, including using incentives to attract effective teachers to work in these schools. Since the program's inception, these schools have improved in both student performance and attendance.[34] Dallas credits the drop in the number of designated low-performing schools (from 43 in the 2013-2014 school year to only four for the upcoming 2018-2019 school year)[35] largely to its evaluation system.[36]



Denver Public Schools (DPS)

Evaluation System: Leading Effective Academic Practice (LEAP)

History: Piloted in 2012-2013[39] and fully implemented in 2013-2014 under Superintendent Tom Boasberg, who continues to lead the district in 2018, although he recently announced that he will resign later this year.

How it works: Teachers earn one of four ratings.

Impact on salary: Maximum: *Undefined. The salary is comprised of base pay and incentive pay for a number of factors (e.g., working in a high-need school or teaching a hard-to-staff subject, receiving a satisfactory evaluation rating). Teachers' future base earnings potential is currently not limited as there is no salary cap on the current step and grade schedule.*

Maximum prior to evaluation system implementation: *$74,218 with a Ph.D. and 12 years of experience.*

REFLECTING ON SYSTEM

Notable features

- Teachers who earn effective ratings are eligible to apply for a range of teacher leadership positions, which allows them to take on a leadership role in their school (and earn up to $5,000 on top of their salaries) without leaving the classroom. These teacher leaders apply the district's Framework for Effective Teaching when coaching and evaluating their peers. To read more about how Denver Public Schools, a Great District for Great Teachers, approaches teacher leadership, visit here.
- The district's research finds that early career teachers tend to improve faster when they are coached on a narrow set of skills and receive direct, bite-sized feedback on those specific skills, rather than trying to develop skills in many areas at once. Using this information, Denver has developed a New Teacher Playbook for early career educators. The playbook prioritizes specific skills and provides guidance intended to help teachers make quick, measurable improvement.[40]
- Principals and their instructional leadership teams can take advantage of optional peer observations. They are paired with an evaluation system specialist who helps school leaders become more consistent in observing teachers and providing high-leverage feedback. In addition, central partners help instructional leadership teams analyze student data to ensure that teachers' growth translates into accelerated student growth.
- Denver requires all observers to complete and pass an initial observation certification process. All observers also engage in an annual school-based calibration session with an evaluation system specialist, which allows the district to identify and support observers whose feedback and ratings may be misaligned.
- In 2015, the district launched incentives for teachers in 30 of its highest-priority schools based on recommendations from teachers. Teachers receive a monthly incentive for working in a highest-priority school and a yearly retention incentive for returning to these schools, with higher-rated teachers earning more. Teachers in the highest-rating category can earn up to $4,000, in addition to other incentives and base pay, for working and staying in a highest-priority school.
- The district views its use of the framework as being pervasive across the district, building the foundation for how it defines excellence in teaching.[41]

> The system is "now integrated into how we think about growth and development, how we look at and leverage renewal decisions, and how we provide support for new teachers and make promotion decisions for leadership positions."
>
> — Nicole Wolden, *Director of Growth and Performance*[52]

Opportunities for improvement

- The district has several mechanisms to identify areas for improvement of the evaluation system and to strengthen those areas. The district regularly engages teachers in these efforts and has created a collaborative committee in partnership with the teachers union.[42] Denver also seeks to build teachers' understanding of and confidence in how the system can be used to help them improve.

Notable changes to system since inception

- Based on feedback from teachers, in some circumstances[43], Denver now allows instructional leaders to incorporate their own judgment (based on a holistic analysis of a teacher's performance) into a teacher's final evaluation rating.
- Informed by feedback from teachers, the district changed the scoring model for how student perception data factor into teacher evaluations. Denver now compares survey data across subsets of teachers, rather than across all teachers in elementary schools or all teachers in secondary schools. For example, student perceptions of elementary homeroom teachers are compared to other elementary homeroom teachers, but not to elementary specials teachers (e.g., art teachers).
- The district created content-specific framework appendices to ensure that the system is adaptable to teachers working with students across various grades, contents, and contexts.

BUILDING BUY-IN

Steps to build buy-in

The district provides training on the evaluation system for all new teachers, focusing on the evaluation system's role in helping teachers grow in their practice.

Surveys of teachers and administrators

On the most recent district survey, more than 80 percent of teachers reported that they felt comfortable having honest conversations about their performance with their coaches and believed that their coaches had an accurate understanding of their effectiveness. Two-thirds of teachers also reported that their experience with the evaluation system helped improve their teaching practice.[44]

Almost 100 percent of school leaders and Senior/Team Leaders responding to the survey reported that they knew what was expected of them to implement LEAP as a growth tool. Nearly 90 percent also reported that the culture in their buildings supports honest assessments of teaching performance.[45]

EVIDENCE OF IMPACT

Provides a more valid measure of teacher quality?

Yes, in 2017-2018, 24 percent of teachers were rated Distinguished, 64 percent Effective, 12 percent Approaching, and less than 1 percent were rated Not Meeting.[46]

Teachers who earn the top two rating categories have significantly higher student growth averages than teachers with lower effectiveness ratings.

Recognizes strong teachers and keeps them in the classroom?

Yes, 91 percent of the highest-rated teachers were retained in the district.[47]

Encourages consistently less effective teachers to leave the classroom?

Yes, 37 percent of teachers in the lowest two rating categories in 2016-2017 were not retained in Denver for the following school year (80 percent of those in the lowest category and 35 of those in the next-lowest category left).[48]

> When an evaluation system has high consequences attached, "how do you take the fear out, and create a system in which both coaching and evaluation live in the same place?"
>
> — Nicole Wolden, *Director of Growth and Performance*[53]

Helps all teachers improve?

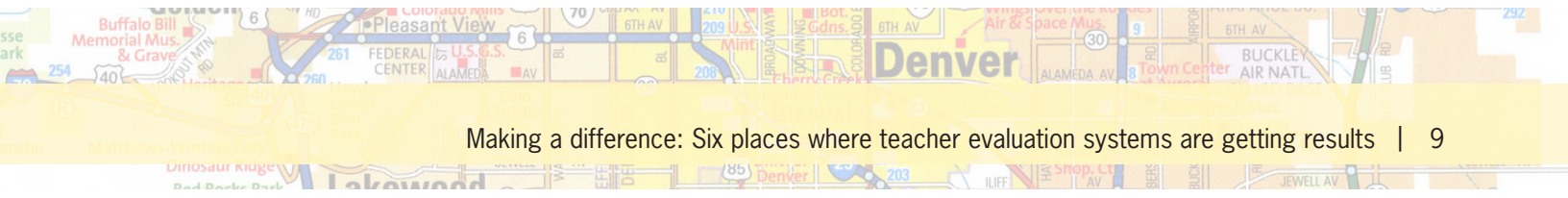
Yes, 92 percent of teachers with evaluation ratings in 2016-2017 and 2017-2018 received either the same rating or improved one rating category, with many teachers receiving the same rating having shown improvement in the underlying components of the framework[49] (although score inflation could also cause these increases).

Recruits more effective new teachers?

Yes, newly hired teachers were on average more effective than exiting teachers in mathematics and were more effective in ELA in most years as well.[50]

Correlates with increased student achievement?

Yes, in the years since Denver has implemented the evaluation system, the district has experienced consistent student growth as measured by state assessments. Denver Public Schools' students have consistently outpaced their classmates statewide in academic growth in English language arts and math.[51] Note that these gains cannot be tied directly to the evaluation system without a controlled study.



District of Columbia Public Schools (DCPS)

Evaluation System: IMPACT

History: Implemented in 2009 under Chancellor Michelle Rhee and continued under Chancellor Kaya Henderson (2010-2016) and then Chancellor Antwan Wilson (2017-2018). Now under Interim Chancellor Amanda Alexander.

How it works: Teachers earn one of five ratings.

Impact on salary: Maximum: *$139,126 (base salary of $114,126 plus up to $25,000 in bonuses based on teaching assignment) with at least nine years of experience.*[54]

REFLECTING ON SYSTEM

Notable features

- The compensation associated with IMPACT is substantial and is the largest of any district or state in this analysis. Teachers can earn up to $25,000 in bonuses annually if they earn the highest of five ratings, teach in a targeted high-poverty school, and meet other criteria.[55] Teachers who consistently earn high ratings can also advance more quickly up the salary schedule.[56] For example, teachers in high-poverty schools who have earned the top rating for six or more years in a row are paid as though they actually had an additional 12 years of experience.[57] So a seventh-year teacher who consistently earned these top ratings is paid as though she is in her 19th year and is moved into the Ph.D. lane of the salary schedule, receiving a base salary increase of nearly $56,000, in addition to any bonuses she earns.[58]

> "Teachers who come to DCPS recognize what IMPACT is, they come knowing that those are the expectations, and they are excited to meet them."
>
> — Betsy Press, *former Deputy Chief of Impact*[70]

- Whereas some districts have struggled with using evaluations for both teacher accountability and improvement, DCPS has created a separate system to help teachers improve, giving them feedback without high stakes attached. The district's professional development program partners teachers with content experts for a cycle of learning, practicing, and receiving feedback. To read more about how DCPS, an Outstanding District for Great Teachers, approaches professional development, visit here.
- DCPS provides an online video library with examples of the district's five "essential practices" of good teaching, in addition to online curricula resources for all subjects.
- To build accuracy in observations, the district requires new principals to go through four hours of online training, in which they watch online videos of instruction and rate teachers to ensure that their ratings are consistent with anchor ratings.

Opportunities for improvement

- A higher percentage of teachers in low-poverty schools receive the highest evaluation rating, while fewer teachers in high-poverty schools receive the highest rating.[59] There is more work to be done in unpacking the root causes of these differences.

Notable changes to system since inception

- DCPS now counts individual value-added measures as up to 35 percent of a teacher's evaluation rating, whereas these used to count for up to 50 percent.
- The district revised its observation rubric so that it is more flexible and streamlined.[60]
- DCPS added student surveys for grades 3 to 12, which now comprise 10 percent of teachers' evaluation ratings.
- The district ended the use of third-party observers (known as Master Educators) in the 2015-2016 school year.

BUILDING BUY-IN

Steps to build buy-in

To make sure teachers understand the evaluation system, DCPS offers IMPACT guidebooks customized for teachers of different grades and subjects. Rather than ask teachers to come to a centralized location, the district provides training in individual schools, which makes it easier for all teachers to learn about the system. The district also provides principals with resources to facilitate their own overview of the evaluation system and introduces all new teachers to IMPACT at the New Teacher Orientation. The IMPACT team is available to answer questions and address teacher and administrator concerns, and responds to inquiries within one business day.

EVIDENCE OF IMPACT

Provides a more valid measure of teacher quality?

Yes, in 2016-2017, 20 percent of teachers received a rating below effective, and 80 percent received one of the top two ratings.[61]

Recognizes strong teachers and keeps them in the classroom?

Yes, DCPS retains 92 percent of its effective and highly effective teachers.[62]

While IMPACT was the top reason given for leaving by 22 percent of the lower-performing teachers who are no longer DCPS teachers, only 6 percent of higher-performing teachers cited it as their top reason for leaving.[63]

> "While the district values stability, we are constantly iterating, reflecting, and always keeping on top of the ways we can improve."
>
> — Betsy Press, former Deputy Chief of Impact[71]

Encourages consistently less effective teachers to leave the classroom?

Yes, since the implementation of IMPACT, low-performing teachers are three times more likely to leave DCPS than high-performing teachers, and account for more than a third of teachers exiting DCPS.[64]

Helps all teachers improve?

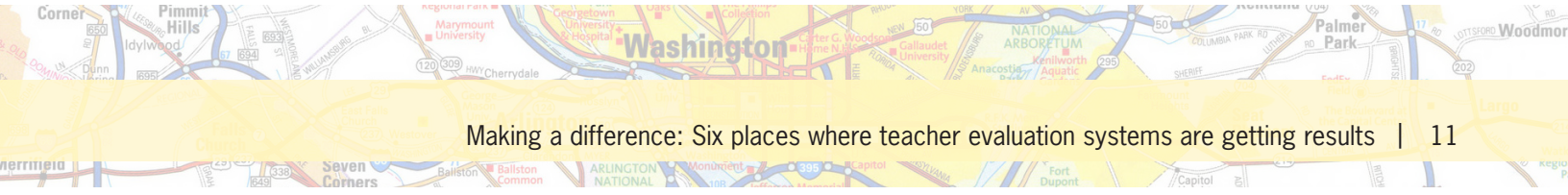
Yes, lower-performing teachers who stayed in the system but were facing threat of dismissal due to a low evaluation rating improved their performance on average. Higher-performing teachers who were near the cut point for a financial incentive also improved.[65]

Recruits more effective new teachers?

Yes, one study found that the incoming teachers increased student achievement by 0.08 SD in math, compared to the teachers who left.[66] Another found that during the first two years of implementation, new hires earned higher ratings on the evaluation system than exiting teachers.[67]

Correlates with increased student achievement?

Yes, since 2009, DCPS has made significant gains on the NAEP assessment in 4th and 8th grade math and reading, although scores declined slightly in math in the most recent NAEP assessment.[68] The white-black student achievement gap has also decreased in 4th grade math and reading since 2009.[69] Note that these gains cannot be tied directly to the evaluation system without a controlled study.



Newark Public Schools (NPS)[72]

Evaluation System: Framework for Effective Teaching

History: Implemented in 2012 under Superintendent Cami Anderson, continued under Superintendent Christopher Cerf through 2018, and continues now under new Superintendent Roger León.

How it works: Newark's evaluation system has four rating categories.

Impact on salary: Maximum: *$100,531 ($95,531 maximum base salary plus up to $5,000 in bonuses), achievable after 18 years.*

Maximum prior to evaluation system implementation: *$103,159, achievable in 29 years with a Ph.D., a masters and 30 credits, or a bachelors with 60 credits.*

REFLECTING ON SYSTEM

Notable features

- The district gave the evaluation system a chance to work. While Newark saw students' achievement decline initially after the new evaluation system was implemented, the district persevered and student achievement rose to the level it had been before and, in English, exceeded previous levels.
- Newark's central office creates monthly reports that serve the dual purposes of providing principals with data on their progress completing evaluations and tracking the integrity of their ratings by showing their schools' teacher ratings distribution compared to other schools' ratings distribution. These reports allow principals, their supervisors, and the central office staff to examine the same data and have conversations around these data.

> The evaluation system is "more than a way to put teachers into buckets; it is for coaching, feedback and growth, and sometimes for exiting or promoting teachers."
>
> — Larisa Shambaugh, *former Chief Talent Officer*[84]

- Newark invested substantial resources when the system was first implemented to train principals on what the language of the evaluation framework meant, what to look for when observing teachers, and how to reliably rate teachers. Now, the district maintains this consistency by holding a brief retraining every summer.
- Newark offers teachers a rebuttal process if they disagree with their evaluation rating.

Opportunities for improvement

- While Newark considers multiple measures in its teacher preparation ratings, it does not use student surveys. Student surveys provide another layer of understanding about a teacher's performance and often align with other measures of teacher effectiveness.

Notable changes to system since inception

- While Newark initially intended to build a video library of best practices, instead they are providing guidance for teachers to videotape themselves and share with their peers in a community of practice.[73]

BUILDING BUY-IN

Steps to build buy-in

Teachers can provide feedback through several panels and advisory committees and have a rebuttal process if they disagree with their final observation rating.[74] Also, much of the professional development is co-developed by teachers.[75]

Surveys of teachers and administrators

According to a 2015 survey, most teachers perceive the evaluation system as valid, accurate, fair, and useful, although support for the differentiated compensation system aligned with the evaluation system is more mixed.[76]

> "Any district that says, 'We nailed that,' is not telling the truth – we could always do more."
>
> — Larisa Shambaugh, *former Chief Talent Officer*[85]

In the same 2015 survey, nearly all school leaders believe that all components of the evaluation system are valid and that the evaluations are an accurate measure of teacher performance.[77]

EVIDENCE OF IMPACT

Provides a more valid measure of teacher quality?

Yes, the most recent available ratings distribution found that 4 percent of teachers were ineffective, 10 percent were partially effective, 76 percent were effective, and 11 percent were highly effective.[78]

Recognizes strong teachers and keeps them in the classroom?

Yes, the retention rate for highly effective teachers is 96 percent in the fifth year of the evaluation system.[79]

Encourages consistently less effective teachers to leave the classroom?

Yes, the attrition rate for ineffective teachers is 49 percent in the fifth year of implementation.[80]

Correlates with increased student achievement?

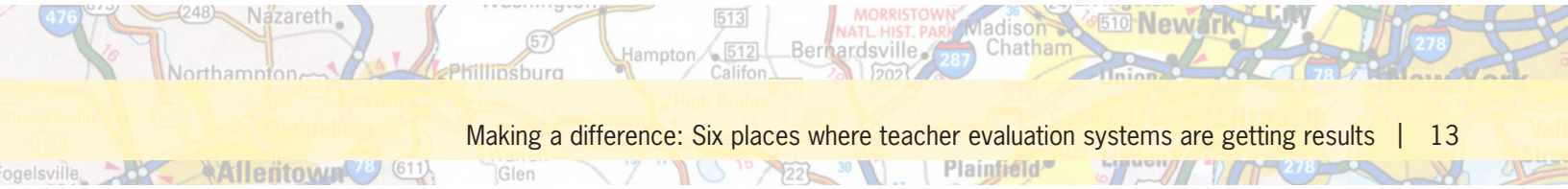
Partially. Since implementing these reforms, Newark Public Schools' students are growing at a faster rate in English language arts compared with their peers across New Jersey, but there is no significant net change in math (after student achievement growth recovered from an initial dip in both subjects).[81] However, much of the gain in English, as well as the growth in math after the initial dip, appears to be driven by the district's decision to move students out of lower- to higher-achieving schools; the teacher evaluation system itself is likely not entirely responsible for the growth.[82]

Newark's schools have made gains in closing the gap between the district's student performance on standardized tests and the state average. Note that these gains cannot be tied directly to the evaluation system without a controlled study.

Other evidence of impact?

Yes, the district has a higher student enrollment now than at any other time in its recent history, a sign of renewed confidence in Newark schools.[83]

> "When ratings go up, we have to ask, are we getting better, or are we inflating? You need to look at other data to keep asking yourself that question in different ways. And in the end, do we see an increase in student learning?"
>
> — Larisa Shambaugh, *former Chief Talent Officer*[86]



New Mexico Public Education Department (NMPED)

Evaluation System: NMTEACH

History: Implemented in the 2013-2014 school year under Governor Susana Martinez and State Education Secretary Hanna Skandera and continues now under Governor Susana Martinez and State Education Secretary Christopher Ruszkowski.

How it works: Teachers earn one of five evaluation ratings.

Impact on salary: New Mexico gives local districts the authority to set pay scales and therefore any connection between teacher salaries and this evaluation system must be decided by each individual district. The state mandates a minimum starting salary, but districts determine the remainder of the schedule.

REFLECTING ON SYSTEM

Notable features

- Requires multiple objective measures: Among all 50 states and the District of Columbia, New Mexico is the only state in which the teacher evaluation system articulates specific percentage requirements for five distinct components. These five components include: student growth, observations, professionalism, student surveys, and teacher absenteeism.
- Includes "teacher absenteeism" as a specific metric: Nationwide, New Mexico is the only state to specifically require that teacher absenteeism be included as part of a teacher's summative evaluation rating. Under this component, comprising five percent of a teacher's overall rating, each New Mexico district submits teacher absences to the New Mexico Public Education Department. Absences due to the Family and Medical Leave Act, bereavement, jury duty, military leave, religious leave, professional development, or coaching are excused. In addition, there is a grace period of up to six days, meaning that if a district reports six days absent for a teacher, the teacher earns 100 percent of the possible points for attendance. If a district reports more than six days absent for a teacher, however, all of that teacher's absences are included in her evaluation.

  Example of possible calculations:

	# of absences reported	Weight	Possible summative points	Summative points earned
Teacher A	0	1	10	10
Teacher B	19	0.05	10	0.5
Teacher C	25	0	10	0
Teacher D	6.5	0.675	10	6.75

- Is fully integrated into New Mexico's teacher preparation, licensure, and support systems:
  - Teacher preparation program accountability: New Mexico requires educator preparation programs to collect and report data on the performance and effectiveness of program graduates, as measured by student growth data.
  - Licensure advancement/renewal: To advance in their licensure, teachers must demonstrate evidence of teacher effectiveness and student learning, which may be demonstrated by earning at least 50 percent of the possible points in the improved student achievement domain of their effectiveness report. To renew their licenses, teachers must demonstrate how they meet the competencies and indicators for their licensure level through their annual evaluation.

> "The impact of including teacher attendance has helped to increase the number of hours that students are served by fully qualified teachers."
>
> — Matthew Montaño, former *Deputy Cabinet Secretary of Teaching and Learning at the New Mexico Public Education Department*[94]

Opportunities for improvement

- New Mexico's definition[87] for "ineffective teacher" under the Every Student Succeeds Act (ESSA) includes not only any teacher who earns an ineffective rating under its teacher evaluation system, but also any teacher who earns a student growth rating in the bottom decile statewide. This definition supports New Mexico's work to ensure that its students — including and particularly the state's low-income students and students of color — have access to teachers with the demonstrated ability to increase student learning and growth. In doing so, it enables New Mexico to better serve some of its most vulnerable students. As it moves forward with its teacher evaluation system implementation, New Mexico should fully integrate its rigorous definition for "ineffective teacher" under the ESSA into its teacher evaluation system. This policy change would help ensure that any teacher who earns student growth ratings in the bottom decile statewide earns an ineffective evaluation rating and has access to the necessary resources and supports to improve her practice.

Notable changes to system since inception

- When New Mexico's evaluation system was first implemented, student growth accounted for 50 percent of a teacher's evaluation rating. As of the 2017-2018 school year, based on recommendations from New Mexico's teachers and leaders, student growth is currently weighted at 35 percent.[88]

> "We are able to distinguish teachers at five levels with a greater understanding of how they're impacting student achievement."
>
> — Matthew Montaño, former *Deputy Cabinet Secretary of Teaching and Learning at the New Mexico Public Education Department*[95]

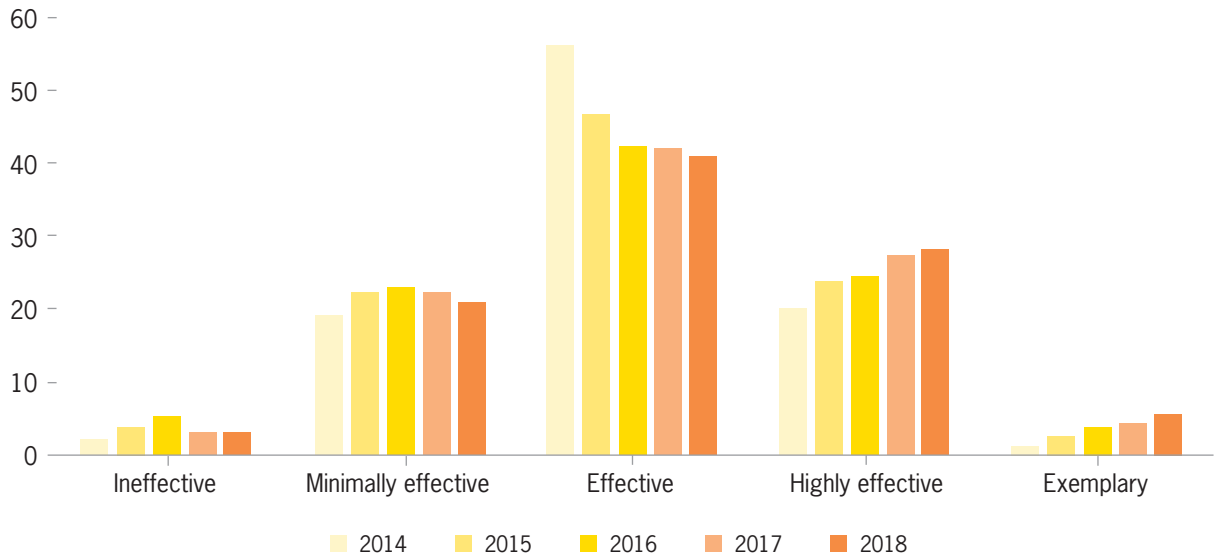
> "We have stuck to the premise that student achievement is not just a part of the evaluation system, but an extremely important part of the evaluation system."
>
> — Matthew Montaño, former *Deputy Cabinet Secretary of Teaching and Learning at the New Mexico Public Education Department*[96]

EVIDENCE OF IMPACT

Provides a more valid measure of teacher quality?

Yes, teachers in New Mexico earn ratings that are substantially more widely distributed across NMTEACH's rating categories than teachers in other states,[89] which enables New Mexico to differentiate the resources and supports available to teachers with differing levels of effectiveness.

New Mexico teacher ratings between 2014-2018[90] :



The percentage of teachers rated below effective among 24 states:[91]

State	Percentage	State	Percentage
New Mexico	28.7	Maryland	2.6
Oregon	11.7	Michigan	2.4
Tennessee	11.4	North Carolina	2.3
Kansas	11.1	Florida	2.2
Arizona	7.0	Idaho	2.2
Louisiana	6.0	Indiana	2.2
Massachusetts	4.9	Georgia	2.0
New York	4.9	Delaware	1.9
Colorado	4.2	Rhode Island	1.7
Connecticut	4.0	New Jersey	1.6
Ohio	3.7	Pennsylvania	1.4
Washington	3.5	Hawaii	0.7

New Mexico's ability to track the percentage of students taught by teachers earning different evaluation ratings is critically important in enabling New Mexico policymakers to make progress in closing its achievement and opportunity gaps. During the 2016-2017 school year, nearly 77 percent of New Mexico's students of color were taught by teachers rated effective, highly effective, or exemplary. In fact, more students of color in New Mexico — both a higher number and a higher percentage — are being taught by highly effective and exemplary teachers than nonminority students.

Helps all teachers improve?

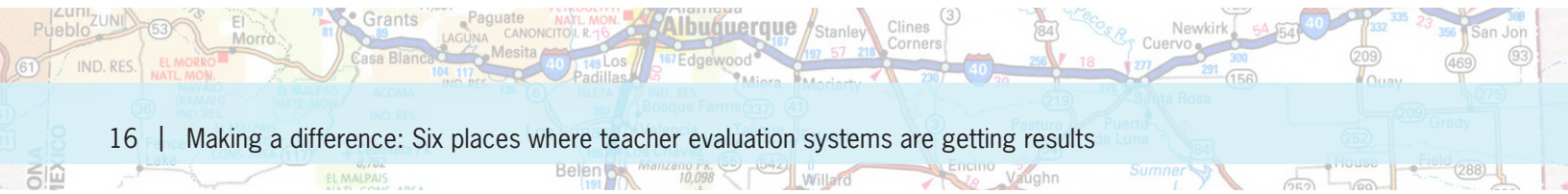
Yes, New Mexico's system allows the state to give teachers precise insight into areas of improvement. The state's teachers receive their evaluation ratings as early as August, and these evaluation ratings are directly linked to professional development opportunities. Between 2015 and 2018, over 1,000 more teachers earned exemplary and highly effective ratings under NMTEACH.[92]

Correlates with increased student achievement?

Yes, between 2015 and 2018, 11,000 more students demonstrated grade-level proficiency in math and 13,000 more students are reading on grade level — with Native American students improving their reading results more than any other group of students, by 8.2 percentage points.[93] Note that these gains cannot be tied directly to the evaluation system without a controlled study.

"Our Teacher of the Year (TOY) is selected with NMTEACH. Therefore, we know our TOY will not have simply won a series of popularity contests. This past year's recipient, Ivonne Orozco, popped on her NMTEACH evaluation, and everyone who has gotten to interact with her can attest to her excellence."

— Chris Eide, *Director of Educator Quality at the New Mexico Public Education Department*[97]



Tennessee Department of Education (TNDOE)

Evaluation System: The Tennessee Educator Acceleration Model (TEAM)

History: Implemented in the 2011-2012 school year under Governor Bill Haslam and Commissioner Kevin Huffman and continues now under Governor Bill Haslam and Commissioner Dr. Candice McQueen.

How it works: Teachers earn one of five ratings.

Impact on salary: Tennessee provides local districts with a minimum salary schedule that sets a base salary with raises earned for each year of experience and type of degree. Distinguishing Tennessee's salary schedule from other states is its requirement that districts must also differentiate teacher compensation based on at least one of the following criteria: additional roles or responsibilities, hard-to-staff schools or subject areas, and performance based on teacher evaluations.

Minimum number of years to achieve highest salary: *After 11 years, Tennessee's state salary schedule provides that teachers with a bachelor's degree will earn $40,595 and those with an advanced degree will earn $45,075.*[98] *These amounts may be supplemented by Tennessee's districts.*

REFLECTING ON SYSTEM

Notable features

- Requires a high number of observations, especially during the first half of the school year. A teacher's prior year performance and license type determine number of observations. Most new teachers are observed six times and most veteran teachers are observed four times — with half of these observations occurring during the first half of the school year. Highly rated teachers are formally observed once but also supported by multiple, short walkthroughs.
- Is fully integrated into Tennessee's teacher preparation, licensure, support, and dismissal systems, with these particularly noteworthy applications:
  - Teacher preparation program accountability: Tennessee requires educator preparation programs to collect and report data on the performance and effectiveness of program graduates, as measured by student growth data.
  - Tenure: A tenured teacher who receives one of the two lowest ratings may be reverted to probationary status until earning a higher rating two years in a row.
  - Dismissal: Unlike 28 states, Tennessee explicitly makes teacher ineffectiveness grounds for dismissal and also requires that teacher evaluation system ratings serve as sufficient evidence of ineffectiveness.[99]

> "Our evaluation system has been a consistent approach of developing the capacity of teachers to improve. The state's TEAM model provides a common language and the vocabulary of continuous improvement. In fact, more than 70 percent of educators believe that the TEAM model has positive impacts for both teachers and students."
>
> — Paul Fleming, *Tennessee's Assistant Commissioner of Teachers and Leaders*[111]

- Independently evaluates the system to determine efficacy. In 2016, Tennessee partnered with Vanderbilt University to form a research alliance (Tennessee Education Research Alliance (TERA)), with the goal of building Tennessee's capacity for continual improvement through independent studies and external research that provide valuable information to state policymakers.[100]

Opportunities for improvement

- Despite the strengths associated with Tennessee's system, it continues to face implementation challenges. Spring 2018 marked the second time in three years that there were significant online testing disruptions of the state standardized assessment, namely log-in delays and connectivity issues, which led the legislature to delay any application of evaluation results to determine employment or compensation.[101]

BUILDING BUY-IN

Steps to build buy-in

In response to feedback from educators during the first year of its evaluation system implementation, Tennessee established a centralized communication strategy to provide teachers and school leaders with the information necessary to successfully implement this system, as well as to learn from educators' implementation experiences. As part of this process, Tennessee met with more than 7,500 educators and considered feedback — surveys, stakeholder meetings, and email communications — from thousands more. In addition, then-Commissioner Huffman visited more than 100 districts to discuss evaluation system implementation with educators.[102]

Surveys of teachers and administrators

According to the recently released 2018 Tennessee Educator Survey, 72 percent of teachers believe the teacher evaluation process used at their school has led to improvements in their teaching. This is up from 38 percent in 2012. Further, 69 percent of teachers believe the teacher evaluation process used at their school has led to improvements in student learning. This is up from 28 percent in 2012. Also, 53 percent of teachers surveyed believe that the feedback received from evaluators was focused more on helping teachers improve, rather than making judgments about performance.[103]

The 2018 Tennessee Educator Survey indicates that 92 percent of administrators feel that they receive useful feedback as a result of the evaluation process, and 91 percent report making changes to their leadership practices as a result of evaluations. Overall, 87 percent of administrators report being satisfied with the evaluation process.[104]

EVIDENCE OF IMPACT

Recognizes strong teachers and keeps them in the classroom?

Yes, Tennessee's teacher retention data demonstrate that teachers earning highly effective ratings are generally retained at a higher rate than less effective teachers. Specifically, in the 2012-2013 school year, more than 90 percent of teachers earning one of the three highest ratings in Tennessee's teacher evaluation system were retained, whereas approximately 80 percent of teachers earning the lowest rating were retained.[105]

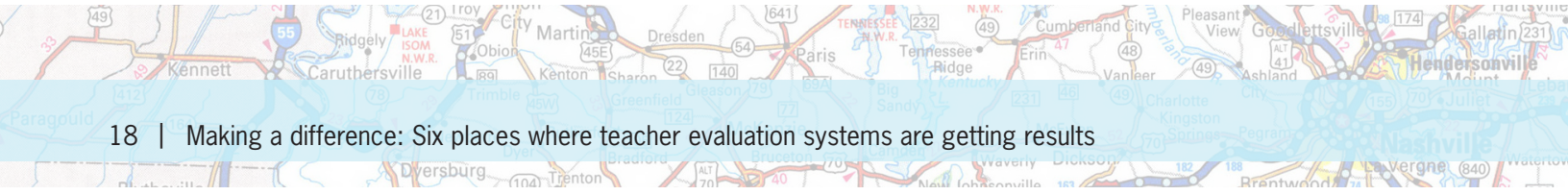
Helps all teachers improve?

Yes, a recent report[106] by TERA found that teacher improvement in Tennessee appears steeper in recent years and coincides with Tennessee's significant investment in efforts to improve teacher professional learning. These investments include Tennessee's development and implementation of its teacher evaluation system in the 2011-2012 school year. Specifically, researchers found that teacher improvement[107] in the years subsequent to the implementation of Tennessee's teacher evaluation system, from 2013 to 2015, was "much more rapid and sustained" than teacher improvement between 2008 and 2010.[108]

> "On organizational theory, it's really important to be clear on the non-negotiables and then provide intentional and intensive supports. Take the time on the front end to develop the rationales so that districts appreciate that this is being done ***with them***, rather than ***to them***."
>
> — Paul Fleming, *Tennessee's Assistant Commissioner of Teachers and Leaders*[112]

Correlates with increased student achievement?

Yes, for the 2016-2017 school year, 55 school districts, equaling more than a third of Tennessee's districts, earned overall Tennessee Value-Added Assessment System (TVAAS) composites representing the highest level of growth. This includes districts with a wide range of academic achievement and student demographics. Further, fewer districts reported the lowest two categories of growth during this same year.[109] These data represent an improvement over the 2014-2015 and 2015-2016 school years, during which 46 and 52 districts, respectively, earned overall TVAAS composites representing the highest level of growth.[110] Note that these gains cannot be tied directly to the evaluation system without a controlled study.



Appendix

Key components of an evaluation system

Multiple measures for final evaluation score: Many factors should be considered in formally evaluating a teacher, with teacher evaluation instruments, including factors that combine both human judgment and objective measures of student learning. The Measures of Effective Teaching (MET) project found that multiple measures — namely classroom observations, student surveys, and measures of student achievement — produce more consistent, stable ratings over the years, as compared to a single measure of effectiveness.[113]

Student surveys: The use of student survey data can provide a rich picture of teacher effectiveness in the classroom and further strengthen an evaluation system's ability to identify teachers' effectiveness.[114] Research finds not only that student input on teacher quality adds value to teacher evaluation systems, but also that teachers often prefer evaluation systems that include student survey data.[115] Further, teacher evaluation systems that include student survey data, which are somewhat correlated with teachers' student growth measures, are stronger, more reliable, and more valid than those that rely solely on administrator reports and observations.[116]

Objective measures of student growth: Teacher evaluation systems should include objective measures of student learning, in addition to factors that rely on informed, subjective judgments of teacher performance. Evidence shows that teachers who increase their students' learning positively influence those students' long-term achievements (e.g., higher likelihood of attending college, earning higher salaries), in addition to benefiting their immediate academic outcomes.[117] Although teachers have many responsibilities, advancing their students' academic achievement is one of every teacher's primary goals.

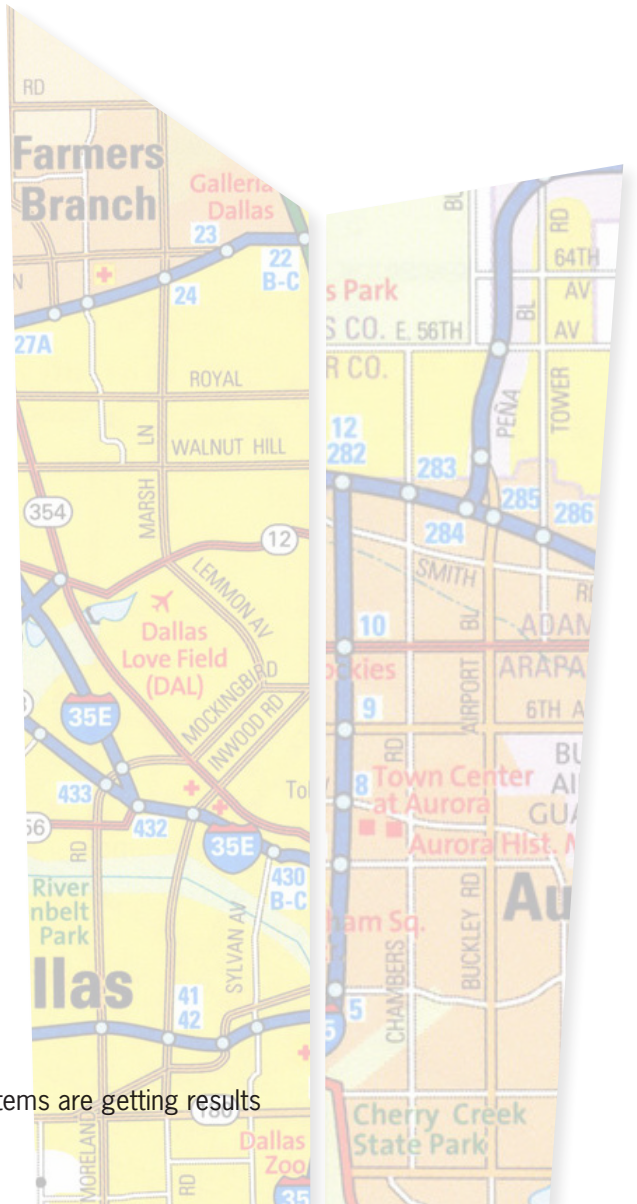
At least three rating categories: Evaluation instruments that differentiate among various levels of teacher performance, rather than those with binary satisfactory/unsatisfactory ratings, further increase the utility and validity of evaluation systems. Traditionally, binary rating systems have offered little meaningful information because virtually all teachers received satisfactory ratings.[118] More rating categories allow for more nuanced distinctions among levels of teacher performance.

Annual observations and evaluations for all teachers: The aim of teacher evaluations is to help all teachers improve, and therefore all teachers need feedback on their performance every year. Observations serve several purposes, including providing actionable feedback to teachers and informing a summative rating that can be used in staffing decisions. Observations are a rich source of information for teachers, giving them useful feedback to improve their practice. Because a teacher's effectiveness can vary from year to year, multiple years of data (e.g., two years of value-added measures (VAM) data) should be used in determining a teacher's summative evaluation rating to increase accuracy.[119]

Professional development tied to evaluations: A primary purpose of teacher evaluations is to identify areas in which a teacher is performing well and areas in which she needs to improve. However, simply naming these areas is likely insufficient. States and districts should also provide professional development that is tailored to a teacher's specific needs.[120] A landmark study of professional development found that, commonly, teachers do not get clear information about how to improve, nor do they feel that their professional development is customized based on their needs.[121]

Written feedback after each observation: Constructive, detailed feedback and coaching is a critical component of teacher development.[122] Requiring feedback in written form (ideally accompanied by a conference between the teacher and observer) can help ensure that observers give deliberate thought to the feedback they impart and provide teachers with guidance that they can revisit as needed.[123] Teachers find regular feedback helpful for improving their instructional practice, and they are more likely to believe that the evaluation system is improving their practice when they receive frequent observations and feedback.[124]

Compensation tied to evaluations: Teachers' salaries can be a tool to recruit and retain teachers and to incentivize or reward certain behaviors.[125] Traditional salary schedules tend to favor years of experience (which generally correspond with increased effectiveness only in a teacher's early years)[126] and level of education (which, except in rare cases like graduate degrees in mathematics, are unrelated to a teacher's effectiveness).[127] Tying compensation to effectiveness, instead of years of experience and attainment of a graduate degree, can create a system that better aligns with a district's values and may be a means of persuading more effective teachers to stay in the district. Additionally, not giving salary increases to ineffective teachers sends an important signal about the significance of teacher effectiveness.



Endnotes

1 Disclaimer: Hannah Putman has family members who work or have worked for the District of Columbia Public Schools central office.

2 Ross, E. et al. (2017). *2017 State Teacher Policy Yearbook National Summary*. Retrieved on September 19, 2018 from https://www.nctq.org/dmsView/NCTQ_2017_State_Teacher_Policy_Yearbook

3 Weisberg, D., et al. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. TNTP. Retrieved August 16, 2018 from https://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness

4 Ibid.

5 National Council on Teacher Quality. (2018). *Teacher Contract Database*. [Data set]. Retrieved August 24, 2018 from https://www.nctq.org/contract-database/home

6 National Council on Teacher Quality. (2017). *State Policy Yearbook Database*. [Data set]. Retrieved August 24, 2018 from https://www.nctq.org/yearbook/home

7 For research supporting these components, see Appendix: Key components of an evaluation system.

8 For nontenured teachers or teachers on a Corrective Action Plan

9 Newark Public Schools requires a conference but not written feedback after an observation. District of Columbia Public Schools requires a conference and written feedback in the form of comments and suggestions against each of the standards in the Essential Practices. However, written feedback isn't required on the same timeline as the post-observation conference (which must happen within 15 days of the observation), but the report does have to be provided by the end of the cycle.

10 Data based on correspondence with Denver Public Schools

11 The weights for Newark's teacher evaluation system are set by state, not district, policy.

12 All districts use observation rubrics the school districts themselves designed.

13 In Newark Public Schools, this section is considered "Teacher practice," in which teachers are evaluated on classroom observations as well as observations of teachers in team meetings and interacting with students, artifacts such as lesson plans and student work, and quantitative data such as surveys.

14 Data based on correspondence with District of Columbia Public Schools

15 This policy is based on Colorado state statute SB 10-191, and is not driven by district policy.

16 Bonuses are limited to teachers who also earn the highest evaluation rating.

17 Bonuses are limited to teachers who also earn the highest evaluation rating.

18 While Dallas does have retention bonuses for teachers who stay in the school district, these are tied to performance rather than increases solely for additional years of service.

19 National Council on Teacher Quality. (2018). *Teacher Contract Database*. [Data set]. Retrieved August 24, 2018 from https://www.nctq.org/contract-database/home

20 Exemplary teachers can earn up to $10,000 annually for working in a designated low-performing school.

21 This requirement was expanded in 2017-18; it used to require that teachers be in the top 25 percent.

22 S. Smith & J. Vega (interview, June 27, 2018)

23 Baeder, J. (2017). Go and see: The key to improving teaching and learning. *ASCD Express, 12*(18). Retrieved July 12, 2018 from http://www.ascd.org/ascd-express/vol12/1218-baeder.aspx

24 Ibid.

25 S. Smith & J. Vega (interview, June 27, 2018)

26 Teacher Excellence Initiative, Dallas Independent School District. (2016). *2016-2017 TEI teacher guidebook: Dallas ISD's guide for defining, supporting and rewarding teacher excellence*. Retrieved June 18, 2018 from https://tei.dallasisd.org/home/resources/

27 Barton, M., & Palladino, D. (2018). *2016-17 Evaluation of Teacher Excellence Initiative (TEI)*. Dallas Independent School District, Department of Evaluation and Assessment. Retrieved June 12, 2018 from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/16-17/finalrpts/EA18-539-2%20Full%202017-18%20TEI%20Report.pdf

28 Ibid.

29 Ibid.

30 Of these, 96 percent continued teaching and 2 percent moved to nonteaching positions in the district. Barton, M., & Palladino, D. (2018). *2016-17 Evaluation of Teacher Excellence Initiative (TEI)*. Dallas Independent School District, Department of Evaluation and Assessment. Retrieved June 12, 2018 from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/16-17/finalrpts/EA18-539-2%20Full%202017-18%20TEI%20Report.pdf

31 Refers to "unsatisfactory" teachers based on several years of evaluation data. Among teachers who were rated "unsatisfactory" in that school year, 43 percent left the district. Barton, M., & Palladino, D. (2018). *2016-17 Evaluation of Teacher Excellence Initiative (TEI)*. Dallas Independent School District, Department of Evaluation and Assessment. Retrieved June 12, 2018 from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/16-17/finalrpts/EA18-539-2%20Full%202017-18%20TEI%20Report.pdf

32 Douglas, N. E., & Barton M. (2017). *Evaluation of the 2015-16 Teacher Excellence Initiative (TEI): Survey results report* (EA17-539-2.2). Dallas Independent School District, Department of Evaluation and Assessment. Retrieved August 16, 2018 from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/15-16/finalrpts/EA17-539_TEI_SurveyAAG.pdf; Barton, M., & Palladino, D. (2018). 2016-17 Evaluation of Teacher Excellence Initiative (TEI). Dallas Independent School District, Department of Evaluation and Assessment. Retrieved June 12, 2018 from https://www.dallasisd.org/cms/lib/TX01001475/Centricity/domain/98/evaluation/16-17/finalrpts/EA18-539-2%20Full%202017-18%20TEI%20Report.pdf

33 Commit! (2017). *Overview of Dallas ISD STAAR Achievement at "Meets" Post-Secondary Standard Across Various Demographics and Subjects 2012-2017*. Dallas, TX: Commit.

34 Palladino, D.K. (2016). *Evaluation of Accelerating Campus Excellence (ACE) 2015-16*. Retrieved July 31, 2018 from https://www.dallasisd.org/site/handlers/filedownload.ashx?moduleinstanceid=66303&dataid=88235&FileName=Evaluation%20of%20Accelerating%20Campus%20Excellence%20ACE%202015-16.pdf; DISD. (2018). ACE accomplishments. Retrieved July 31, 2018 from https://www.dallasisd.org/Page/46766

35 Smith, C. (2018, June 22). How many Dallas ISD schools failed to meet state accountability? DISD leader reveals an amazing number. *Dallas News*. Retrieved June 28, 2018 from https://www.dallasnews.com/news/education/2018/06/22/many-dallas-isd-schools-failed-meet-state-accountability-disd-leader-reveals-amazing-number

36 S. Smith & J. Vega (interview, June 27, 2018).

37 Ibid.

38 Ibid.

39 LEAP was piloted as part of the Measures of Effective Teaching Project.

40 N. Wolden (personal communication, 2018).

41 Ibid.

42 N. Wolden (interview, June 19, 2018).

43 Leaders may use their judgement for teachers whose calculated Professional Practice ratings fall between performance categories or whose Professional Practice and Student Growth ratings were different.

44 Wolden, N. (2018). LEAP Pre-Reading. Shared via personal communication.

45 N. Wolden (personal communication, 2018).

46 Ibid.

47 Education Resource Strategies. (2017). *Denver Public Schools: Leveraging System Transformation to Improve Student Results*. Retrieved July 12, 2018 from https://www.slideshare.net/ERSslides/ers-system-2020-analysis-of-denver-public-schools-march-2017

48 N. Wolden (personal communication, 2018).

49 Ibid.

50 Atteberry, A., Briggs, D. C., LaCour, S., & Bibilos, C. (2015). *Year 2 Denver ProComp evaluation report: Teacher retention and variability in bonus pay, 2001-02 through 2013-14*. Colorado Assessment Design Research and Evaluation (CADRE) Center.

51 Denver Public Schools. (2018). DPS Students Again Outpace State in Academic Growth on 2018 CMAS. Retrieved October 1, 2018 from https://www.dpsk12.org/dps-students-again-outpace-state-in-academic-growth-on-2018-cmas/

52 N. Wolden (interview, June 19, 2018).

53 Ibid.

54 Maximum salary is based on FY 2018 ET 15 Salary Schedule. Retrieved August 8, 2018 from https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/ET-15%20FY%2018%20Pay%20Schedule.pdf. Maximum salary requires 21 years of experience and a Ph.D., but through the Leadership Initiative For Teachers (LIFT) career ladder, a teacher in a high-poverty school can receive an additional 12 service credits and can be bumped to the master's and then the Ph.D. lane, thus reaching the maximum salary in nine years.

55 District of Columbia Public Schools. (2017). *IMPACT: The District of Columbia Public Schools Effectiveness Assessment System for School-Based Personnel: 2017-2018*. Retrieved July 12, 2018 from https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/1%20%20%20Teachers%20Grades%204%20with%20Individual%20Value-Added%20Student%20Achievement%20and%20Student%20Survey%20Data.pdf

56 District of Columbia Public Schools. (2017). *LIFT: Leadership Initiative For Teachers: 2017-2018*. Retrieved July 12, 2018 from https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/2017-18%20LIFT%20guidebook%20-%20FINAL.pdf

57 Teachers earn two years of service credits when moving into the Advanced stage, five for moving into the Distinguished stage (plus movement to master's lane if not already there), and five for moving into the Expert stage (plus movement into the Ph.D. lane if not already there), all if the teacher is in a high-poverty school.

58 District of Columbia Public Schools. (2017). *IMPACT: The District of Columbia Public Schools Effectiveness Assessment System School-Based Personnel: 2017-2018*. Retrieved July 12, 2018 from https://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/1%20%20%20Teachers%20Grades%204%20with%20Individual%20Value-Added%20Student%20Achievement%20and%20Student%20Survey%20Data.pdf

59 Chandler, M. A. (2015, February 24). IMPACT scores ward-by-ward in D.C. *The Washington Post*. Retrieved July 12, 2018 from https://www.washingtonpost.com/news/local/wp/2015/02/24/impact-scores-ward-by-ward-in-d-c/?utm_term=.f69a4e47b0b5; Toch, T. (2018). *A policymaker's playbook: Transforming public schools in the nation's capital*. Washington, DC: FutureEd Georgetown. Retrieved July 11, 2018 from https://www.future-ed.org/a-policymakers-playbook-for-transforming-teaching/

60 District of Columbia Public Schools. (2010). *IMPACT: The District of Columbia Public Schools Effectiveness Assessment System for School-Based Personnel: 2010-2011*. Retrieved July 12, 2018 from https://dcps.dc.gov/sites/default/files/https://www.nctq.org/dmsView/Impact_The_District_of_Columbia_Public_Schools_Effectiveness_Assessment_DCPS_Report

61 C. Lewis (personal communication, 2018).

62 Pennington, K., & Brand, A. (2018). *Retaining high performers: Insights from DC Public Schools' teacher exit survey*. Washington, DC: Bellwether Education. Retrieved June 15, 2018 from https://bellwethereducation.org/publication/retaining-high-performers-insights-dc-public-schools-teacher-exit-survey

63 Ibid.

64 Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis, 39*(1), 54-76.

65 Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management, 34*(2), 267-297.

66 Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis, 39*(1), 54-76.

67 Walsh, E., & Dotter, D. (2014). *Longitudinal analysis of the effectiveness of DCPS teachers (No. 40185.533)*. Mathematica Policy Research.

68 District of Columbia Public Schools. (2018, April 20). *DC Public Schools continues steady growth on NAEP for ninth year*. Retrieved July 12, 2018 from https://dcps.dc.gov/release/dc-public-schools-continues-steady-growth-naep-ninth-year

69 The Nation's Report Card. (2018). District of Columbia Student Groups and Gaps Data. Retrieved July 12, 2018 from https://www.nationsreportcard.gov/profiles/stateprofile/overview/DC?cti=PgTab_GapComparisons&chort=1&sub=RED&sj=DC&fs=Grade&st=MN&year=2009R3&sg=Race/Ethnicity:_White_vs._Black&sgv=Difference&ts=Cross-Year&tss=2017R3-2009R3&sfj=NP

70 B. Press (interview, June 15, 2018).

71 Ibid.

72 This case study was reviewed for accuracy by former Chief Talent Officer Larisa Shambaugh. Current district leadership did not respond to requests to review the information before publication.

73 L. Shambaugh (interview, June 12, 2018).

74 Ibid.

75 L. Shambaugh (interview, June 12, 2018).

76 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report. *American Institutes for Research*.

77 Ibid.

78 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). *Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research. 2017-18 data comes from personal communication with Larisa Shambaugh, Newark Chief Talent Officer. June 11, 2018.

79 Fulbeck, E., Citkowicz, M., Hester, C., Manzeske, D., Yisak, M., & Eisner, R. (2016). *Newark Public Schools and Newark Teachers Union teacher contract evaluation: Year 1 report*. American Institutes for Research. 2017-18 data comes from personal communication with Larisa Shambaugh, Newark Chief Talent Officer. June 11, 2018.

80 Ibid.

81 Chin, M., Kane, T. J., Kozakowski, W., Schueler, B. E., & Staiger, D. O. (2017, October). *Assessing the impact of the Newark education reforms: The role of within-school improvement vs. between-school shifts in enrollment*. Center for Education Policy Research. Retrieved from https://cepr.harvard.edu/files/cepr/files/newark_ed_reform_report_synopsis.pdf; Newark Public Schools (2016, November). Newark annual report. Retrieved from http://www.nj.gov/education/sboe/meetings/2016/November/public/Newark%20Annual%20Report.pdf

82 Ibid.

83 Margolis, J. (2017, October). *Moving up: Progress in Newark's schools from 2010 to 2017*. MarGrady Research. Retrieved from http://margrady.com/wp-content/uploads/2017/10/Moving-Up-Progress-in-Newarks-Schools.pdf

84 L. Shambaugh (interview, June 12, 2018).

85 Ibid.

86 Ibid.

87 https://www2.ed.gov/admins/lead/account/stateplan17/nmcsa2017.pdf

88 Kim Burgess, "PED alters teacher evaluation system after pushback," *Albuquerque Journal* (April 2, 2017).

89 Kraft, M. A., & Gilmour, A. F. (2017, July 17). Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher. *Educational Researcher, 46*(5), pp. 234-249.

90 New Mexico Public Education Department. (2018). Elevating the Teaching Profession in New Mexico: 2018 NMTEACH Results, Statewide Overview. Retrieved on October 2, 2018 from https://webnew.ped.state.nm.us/

91 Kraft and Gilmour's target sample included 38 states that had either piloted or fully implemented a new teacher evaluation system by the 2014-2015 school year. Their search produced data on the distribution of teacher effectiveness for 24 states.

92 New Mexico Public Education Department. (2018). New Mexico has more Exemplary and Highly Effective teachers than ever before. Retrieved October 2, 2018 from https://webnew.ped.state.nm.us/news/new-mexico-has-more-exemplary-and-highly-effective-teachers-than-ever-before/

93 Albuquerque Journal Editorial Board, "Editorial: Student gains a strong reason to keep PARCC," *Albuquerque Journal* (July 14, 2018).

94 M. Montaño (phone interview, June 15, 2018).

95 Ibid.

96 Ibid.

97 C. Eide (email correspondence, July 20, 2018)

98 Tennessee Department of Education. (2017). *State Salary Schedule* (Tenn. Code Ann. 49-3-306). Retrieved July 12, 2018 from https://www.tn.gov/content/dam/tn/education/cpm/loc_fin/loc_fin_state_min_salary_sched_2017-18.pdf

99 https://www.nctq.org/yearbook/national/Dismissal-79

100 Tennessee Education Research Alliance: vu.edu/TNEdResearchAlliance

101 Herold, B. (2018, May 8). Online state testing in 2018: Mostly smooth, with one glaring exception. *Education Week*.

102 Tennessee Department of Education: First to the Top. (2012, July). *Teacher evaluation in Tennessee: A report on year 1 implementation*. Retrieved September 7, 2018 from: https://www.tn.gov/content/dam/tn/education/reports/rpt_teacher_evaluation_year_1.pdf

103 2018 Tennessee Educator Survey: Statewide Results for Main Survey, http://educatorsurvey.tnk12.gov/#1/all-districts/all-schools/0. Bailey, I. (n.d.). *Takeaways from the 2017 Tennessee Educator Survey*. Educator Insights, Tennessee Department of Education. Retrieved July 12, 2018 from https://www.tn.gov/education/data/tvaas.html

104 2018 Tennessee Educator Survey: Statewide Results for Main Survey. Retrieved September 7, 2018 from http://educatorsurvey.tnk12.gov/#1/all-districts/all-schools/0.

105 Office of Research and Policy Policy Brief. (2014, May). *Teacher retention in Tennessee: Are we keeping our best teachers?* Retrieved September 7, 2018 from: https://www.tn.gov/content/dam/tn/education/reports/rpt_teacher_retention.pdf

106 *Exploring Teacher Improvement in Tennessee: A Brief on Reimagining State Support for Professional Learning*. Retrieved September 7, 2018 from https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/Exploring_Teacher_Improvement.pdf

107 To measure teacher improvement over time, Papay and Laski estimated how individual teachers improved over the course of their careers (i.e., how the effectiveness of a 10-year veteran teacher compares to *her own* effectiveness as a novice). Researchers examined student test scores as well as teacher evaluation ratings — both observation scores and teacher value-added ratings — to measure teacher effectiveness over time.

108 *Exploring Teacher Improvement in Tennessee: A Brief on Reimagining State Support for Professional Learning*. Retrieved September 7, 2018 from https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/Exploring_Teacher_Improvement.pdf

109 Tennessee Department of Education. (2017). *State salary schedule (Tenn. Code Ann. 49-3-306)*. Retrieved July 12, 2018 from https://www.tn.gov/education/data/tvaas.html

110 Tennessee Department of Education. Data Downloads and Requests. Additional Data. TVAASComposites. District Wide TVAAS. 2015-16 and 2014-15. Retrieved September 7, 2018 from: https://www.tn.gov/education/data/data-downloads.html

111 P. Fleming (phone interview, June 1, 2018).

112 P. Fleming (phone interview, June 1, 2018).

113 Cantrell, S. & Kane, T. J. (2013, January). *Ensuring fair and reliable measures of effective teaching*. The Bill and Melinda Gates Foundation.

114 Ibid.

115 Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education, 14*(2), 135-153; Peterson, K. D. (2004). Research on school teacher evaluation. *NASSP Bulletin, 88*(639), 60-79; Stronge, J., & Ostrander, L. (1997). Client surveys in teacher evaluation. In J. H. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 129-161). Thousand Oaks, CA: Corwin Press.

116 Peterson, K. D., & Stevens, D. (1988). Student reports for school teacher evaluation. *Journal of Personnel Evaluation in Education, 1*, 259-267; Stronge, J., & Ostrander, L. (1997). Client surveys in teacher evaluation. In J. H. Stronge (Ed.), Evaluating teaching: A guide to current thinking and best practice (pp. 129-161). Thousand Oaks, CA: Corwin Press.

117 Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633-79. Jackson, C. K. (2012). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina (Working Paper No. 18624). Cambridge, MA: National Bureau of Economic

Research. Retrieved from http://www.nber.org/papers/w18624

118 Weisberg, D., et al. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New Teacher Project. Retrieved August 16 2018 from http://files.eric.ed.gov/fulltext/ED515656.pdf

119 Cantrell, S., & Kane, T. J. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study. *MET Project Research Paper*. Retrieved August 8, 2018 from http://k12education.gatesfoundation.org/resource/ensuring-fair-and-reliable-measures-of-effective-teaching-culminating-findings-from-the-met-projects-three-year-study/

120 For a summary of research on the importance of using compensation as an incentive or report, see NCTQ. (2017). Professional support: The district's professional development is high quality and tailored to teachers' needs. *Great Districts for Great Teachers*. Retrieved June 29, 2018 from http://www.greatdistricts.org/district/criteria.do?detail=true#slide-2-1; and NCTQ. (2017). Professional support: Teachers receive feedback and coaching to help them improve their performance. *Great Districts for Great Teachers*. Retrieved June 29, 2018 from http://www.greatdistricts.org/district/criteria.do?detail=true#slide-2-4

121 Jacob, A., & McGovern, K. (2015). *The Mirage: Confronting the hard truth about our quest for teacher development*. TNTP. Retrieved June 29, 2018 from https://tntp.org/assets/documents/TNTP-Mirage_2015.pdf

122 Research establishes that teachers value and respond to high-quality feedback. For instance, one teacher survey found that the preference for frequent observations and detailed feedback (as opposed to few observations and general feedback) stood at 75 percent for millennial teachers, 70 percent for Generation X teachers, and 59 percent for baby boomers. Coggshall, J. G., Behrstock-Sherratt, E., & Drill, K. (2011). Workplaces that support high-performing teaching and learning: Insights from Generation Y teachers. *American Institutes for Research and the American Federation of Teachers*. Retrieved August 16 2018 from http://files.eric.ed.gov/fulltext/ED518392.pdf. See also Greenberg, M. (2015). Teachers want better feedback. *Education Week, 35*(1), p.48. Retrieved August 16 2018 from http://www.edweek.org/ew/articles/2015/08/19/teachers-want-better-feedback.html. Teacher response to feedback includes improvement in performance (Taylor & Tyler 2011) and greater longevity in the profession (Jacob, 2012). Taylor, E. S., & Tyler, J. H. (2011). The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers. *National Bureau of Economic Research*. Retrieved August 16, 2018 from http://www.nber.org/papers/w16877.pdf. Jacob, A. (2012). The irreplaceables: Understanding the real retention crisis in America's urban schools. New York City: TNTP. Retrieved August 16, 2018 from https://tntp.org/publications/view/the-irreplaceables-understanding-the-real-retention-crisis

123 Wayne, A. J., et al. (2016). *Early implementation findings from a study of teacher and principal performance measurement and feedback: Year 1 Report (NCEE 2017-4004)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved Auguest 16, 2018 from https://ies.ed.gov/ncee/pubs/20174004/pdf/20174004.pdf. Duchaine, E. L., Jolivette, K., & Fredrick, L. D. (2011). The effect of teacher coaching with performance feedback on behavior-specific praise in inclusion classrooms. *Education and Treatment of Children, 34*(2), pp. 209-227.

124 Prado Tuma, A., Hamilton, L. S., & Tsai, T. (2018). *A nationwide look at teacher perceptions of feedback and evaluation systems: Findings from the American Teacher Panel*. Santa Monica, CA: RAND Corporation. Retrieved July 12, 2018 from https://www.rand.org/pubs/research_reports/RR2558.html

125 For a summary of research on the importance of tailoring professional development to teachers' individual needs, see NCTQ. (2017). Professional compensation structure: Salaries reward teachers for doing a great job. Great Districts for Great Teachers. Retrieved June 29, 2018 from http://www.greatdistricts.org/district/criteria.do?detail=true#slide-1-2

126 Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics, 100*, 61-78. Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. Journal of Public Economics, 95(7-8), 798-812. Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. Journal of Public Economics, 130, 105-119.

127 A number of studies have concluded that higher degrees do not lead to greater teacher effectiveness. For example, see, Bastian, K. C. (2018). A degree above? The value-added estimates and evaluation ratings of teachers with a graduate degree. *Education Finance and Policy* (Just Accepted), 1-46. For a summary of additional research, see Walsh, K., & Tracy, C. O. (2004). Increasing the odds: How good policies can yield better teachers. National Council on Teacher Quality. Retrieved June 29, 2018 from https://www.nctq.org/dmsView/Increasing_the_Odds_How_Good_Policies_Can_Yield_Better_Teachers_NCTQ_Report

National Council on Teacher Quality

1440 G Street, NW, Suite 8207
Washington, D.C. 20005
Tel: 202 393-0020
Web: www.nctq.org