



Date: August 26, 2020

Prepared By: Canada

Purpose: Review of Teacher Evaluation Task Force Recommendations

Witness: Katya Danielle Gothie, Director, Educator, Growth and Development Bureau, PED; Gwen Perea Warniment, Ph.D., Deputy Secretary for Teaching, Learning and Assessment, PED; Linda Darling Hammond, Ed.L.D., President and Chief Executive Officer, Learning Policy Institute

Expected Outcome: Understand the new educator evaluation system, including potential student learning metrics that can be included in the model.

New Mexico Teacher Evaluation System and Task Force Recommendations

Federal Policy Influence on Historically Locally-Controlled Teacher Evaluations

Historically, school districts held the responsibility of evaluating teachers. Measures to capture teacher performance and the purpose of the evaluation tool was also left up to local decision-making. Evaluation tools primarily relied on principal observations of their teachers. States began to modify teacher evaluations when federal law started to influence evaluations with a goal to increase teacher quality. Over the past decade, teacher evaluation policies across the nation have changed significantly with federal influence and the corresponding pushback from educators concerned that teacher evaluation strategies did not accurately reflect educators practice.

In response to federal incentives, many states changed their laws to make their application for the federal Race to the Top grant program more competitive and to receive waivers from requirements of the federal No Child Left Behind Act. A requirement of these federal reform efforts was to include student achievement data in teacher evaluation systems. In response, starting in 2009, 15 states began to require objective measures of student growth, a shift from the traditional approach to teacher evaluations. By 2015, 43 states included student academic measures in their evaluation systems. Although the federal government continues to emphasize the use of state teacher evaluation systems through the Every Student Succeeds Act, the act changed the federal role in teacher evaluation by prohibiting the U.S. secretary of education from forcing states to set-up specific teacher evaluation policies. While states have largely kept the improved evaluation systems implemented pre-2015, at least 30 of these systems have been changed significantly in an attempt to provide more balanced evaluations of teachers.

Research shows “fair and balanced” teacher evaluations require multiple measures of teacher performance and should include objective measures.

States started using student state standardized test scores in calculating student growth measures. Currently, only 24 states require this data. Other tools used to measure student growth include school district assessments, student portfolios, and student learning objectives.

New Mexico’s Approach to Teacher Evaluation From 2011 to Present

In 2011, newly elected Governor Martinez issued an executive order to create the Effective Teaching Task Force to identify, recruit, reward, and retain high-performing teachers in New Mexico. According to the order, the new evaluation system should help identify the most effective teachers, provide targeted assistance for those in need, and be used for personnel decisions. The executive order required

At the time, the task force recommended the use of a value-added model of data analysis with a goal to reliably capture student achievement. The model would be measured through the NM Standards Based Assessment. For grades and subjects not covered in the test, the task force recommended other PED-approved local assessments should be used as a proxy.

Teacher evaluation ratings traditionally rated teachers as satisfactory or unsatisfactory. Between 2011 and 2017, states with more than two categories in teacher evaluation scores increased from 17 states to 44 states. New Mexico's system in 2012 rated teachers in one of five categories. PED's new guidance on the system decreases the scoring categories to four.

A law enacted during the 2019 legislative session explains teachers' personal leave and up to 10 sick days will not affect teachers' annual evaluations, unless the leave is used in a manner inconsistent with local school board policies.

at least 50 percent of the new evaluation score to include student achievement as a way to measure teacher effectiveness. The executive order further directed the task force to determine the remaining elements of the evaluation based on best practices of effective teaching and suggest how much the other elements should be weighted. According to the executive order, the task force needed to explore evaluation models, a performance pay structure, value-added determinations, and tenure.

Based on these requirements, the task force recommended a new teacher evaluation system where 50 percent of a teacher's evaluation score would be based on student test scores, 25 percent based on observations, and 25 percent through locally adopted Public Education Department (PED)-approved multiple measures. During the 2012 legislative session, a bill to create the system permanently in law was never passed. Following the legislative session, PED created the system through regulation. Teacher unions in the state sued the department claiming an overstep in authority, which the court denied. During the 2013-2014 school year, the department implemented the new system in line with the weighted elements prescribed by the task force. Within a span of a few years, multiple changes were made to the system, including during the 2015-2016 school year expanding the evaluation to include scoring based on surveys and teacher attendance. A Brown University study found that New Mexico rated nearly 29 percent of teachers in the bottom two categories, below effective during the 2015-2016 school year, while most states placed fewer than 4 percent of teachers in that category. Nationally, New Mexico's results garnered attention and assertions that the system was one of the toughest in the nation. Lawsuits from multiple teacher unions in the state stated the evaluation was scoring teachers unfairly. Up until 2018, multiple modifications to the system occurred, such as decreasing the weight of the student-growth component and allowing teachers more flexibility to use their accrued leave. In 2019, PED under a new administration retroactively amended the summative reports from the 2018-2019 school year to exclude student achievement growth and teacher attendance. A memo to school leaders specified the reports should be used to create the teacher's professional development plan for the 2019-2020 school year. Teacher evaluations were canceled later in the year due to the physical closures of schools from the COVID-19 public health emergency making it unclear if professional development plans were canceled as well.

New Approach to Teacher Evaluation and Task Force Recommendations

Beginning in 2019, newly elected Governor Lujan Grisham tasked the department to engage stakeholders across the state to develop a teacher evaluation system that provides support for educators to improve their practice. Forty-six members of the New Mexico Teacher Evaluation Task Force were appointed from a pool of nominated state-wide education stakeholders including educator preparation program staff, educational leaders, teachers, unions, parents, school administrators, and advocates. Meetings of task force members started in August 2019. Task force members were charged to review stakeholder data to determine a transition



evaluation system for the 2019-2020 school year and create a new comprehensive teacher evaluation system that includes elements beyond observation, such as surveys and growth measures. PED recently released the task force's seven recommendations that will advise the department on elements of the new system. Additionally, PED started issuing guidance for the new system, requiring school districts "test" the system during the 2020-2021 school year.

Guidance from the department on the new system shows the system will no longer be focused on evaluating teachers directly through student test scores and teacher attendance. The new evaluation system specifies the system will be based on three elements: a professional development plan, multiple observations, and surveys. As PED implements the new evaluation system, the department must provide clarity on some outstanding elements of the system. For example, PED's guidance states the focus of the system is to guide educators through a continuous cycle of improvement, but does not address how it will be used to make personnel decisions, as required by rule and law. The department must prioritize updating rule to reflect all of the current system requirements and expectations. Additionally, although PED has provided some guidance on "testing" the new system this year, details are still lacking on how much weight each element will contribute to an educator's overall evaluation score.

Leading up to the creation of the Teacher Evaluation Task Force, a total of 23 regional stakeholder meetings were held across the state with parents, students, teachers, administrators, and community members. Meetings were held in Albuquerque, Artesia, Bernalillo, Clovis, Deming, Espanola, Farmington, Gallup, Hobbs, Las Cruces, Las Vegas, Ruidoso, and Santa Fe.

Recommendation One: Purpose of the System

The task force wants to expand the scope of a newly designed system to be inclusive of all educators, not just teachers. The task force recommends PED design the new system to improve educator and student learning, growth, and well-being. According to the task force, the system also must support meaningful, actionable feedback and professional self-reflection. The system should aim to strengthen a learning culture through communication, collaboration, continuous improvement, and shared ownership. These values align with what was identified during stakeholder engagement. Stakeholders wanted a new system that supports educator growth and moves away from being punitive. Stakeholders also want the new system to support educators as they take risks to improve student learning, give them opportunity to improve their practice, and encourage peer and admin collaboration.

Using the Teacher Evaluation Tool. Through national education reform, federal requirements also influenced how teacher evaluations were used to make personnel decisions. Nationally the tool changed to inform decisions about tenure and promotion, advanced certification, and dismissal. Previously, teacher evaluations were used to guide additional support for personnel through professional development. The state School Personnel Act requires PED to create a "highly objective uniform statewide standard of evaluation" for the annual performance evaluation of licensed school employees and school administrators can use the tool to make personnel decisions. Currently in statute, if a level 2 or level 3-A teacher's performance evaluation shows the individual is showing a less than satisfactory performance and competency, the school principal may require the teacher to participate in peer intervention, which may include mentoring. If performance is not improved in the time period determined by the school principal, the peer interveners may recommend termination of the teacher. For a level 3-A teacher, according to PED

Charlotte Danielson created the Framework for Teaching over two decades ago as a research-based tool that describes the components of effective instruction. The tool divides the activity of teaching into 22 components clustered into four domains of teaching responsibility. The four domains are: (1) planning and preparation, (2) classroom environment, (3) instruction, and (4) professional responsibilities. The Danielson Group claims the tool was developed to define great teaching and provide a comprehensive approach to teaching and learning from pre-service teacher preparation through teacher leadership. This tool is aligned with the Common Core State Standards.

rule, before a termination is recommended, the teacher can be issued a level 2 license as a mechanism of suspension until they demonstrate performance and competencies for the level 3-A license. In December 2015 an injunction was granted from The 1st Judicial District Court preventing consequential actions based on the evaluations, which at the time relied heavily on student achievement data. The tool could not be used to make employment, advancement, or licensure decisions until the system was reliable, valid, fair, and uniform. A trial on the merits was never scheduled and the case was assigned to a new judge. However, in April 2019 according to one of the plaintiffs, they are no longer pursuing the lawsuit due to the department no longer supporting practices opposed by the teacher unions. Although current law and rule allow personnel decisions to be made using the teacher evaluation tool, the task force recommendations and PED's new system does not address this aspect of the tool. The agreed upon focus of the new evaluation system from both the task force and PED is on developing teachers and providing them with feedback on their practice.

Recommendation Two: System Design

PED announced on July 29, 2020, that the new educator evaluation system, Elevate New Mexico, developed during the 2019-2020 school year will be tested by all New Mexico educators during the 2020-2021 school year, with formal implementation of the new system happening the following year. Components of the evaluation will include a professional development plan, observations, and surveys. These elements are in line with what was recommended by the task force, but PED did not specify in their most recent guidance how each element will be weighted within the final evaluation score where educators are identified as not demonstrating, developing, applying and innovating. According to the task force, half of the score should be based off of observations and feedback using the Charlotte Danielson's Framework rubric, which is currently used. PED specified the new system would include informal observations of educators each quarter and one final annual observation related to creating an environment and teaching for learning based on the rubric. The task force recommends the other half of the system be based on an educator-created professional development plan. Administrators are encouraged by PED to use the tool this year to assist educators on focusing on teaching in a remote or hybrid environment, familiarizing themselves with technology, and supporting the social and emotional wellbeing of students. The task force recommends evidence of student learning be used to inform the professional development plan through student survey and student performance data, but recommends isolated student achievement data does not contribute to an educator's overall performance assessment. PED's guidance on the new system requires educators to base their professional development plan on student-centered goals using data and evidence to inform the plan. Additionally, the task force advocates for mentorship to be used formatively to support and inform educator professional performance and assessment. Although the new evaluation system will be tested this year, PED needs to specify how each element will contribute to the educator's overall evaluation score.

Integrated Teaching and Learning System The task force suggested elements that are currently required by law should continue to be included in and interact with the evaluation system like observations, professional development plans, and mentorship. PED should ensure all elements that interact with the evaluation system continue to complement one another so educators are clear on expectations and the standards they are held to throughout their careers. According to Linda Darling Hammond, Ed.L.D., an expert in teacher preparation, creating a coherent system from preparation through practice will greatly improve the capacity of the teacher workforce. Teacher evaluation must be embedded within the entire system from preparation and induction programs through daily professional practice opportunities. In 2003, the New Mexico Legislature adopted multiple education reforms which included a career ladder for teachers and school administrators and aligned specific requirements with the federal No Child Left Behind Act. The three-tiered licensure system put in place in 2003 requires beginning teachers to undergo a formal mentorship program and be evaluated annually. To advance from a level 1 license to a level 2 or level 3-A teaching license, a superintendent must verify competency through the evaluation tool and a teacher must submit teaching evidence through a professional development dossier, a portfolio process modeled after the National Board for Professional Teaching Standards. In line with best practice, the 2003 professional framework was created for education licensing exams, teacher evaluation, mentorship, ongoing professional development, the professional development dossier, and the professional development plan to work together and complement each other based on a shared understanding of New Mexico's teacher licensure competencies and the Common Core State Standards. PED's implementation of the new professional development plans are not based on either set of standards, although current rule requires the professional development plan be based on the nine New Mexico teacher competencies. PED will need to change rule to reflect this difference or clarify on what standards the professional development goal will be based on. For the three-tiered licensure system to remain intact, measures and expectations behind the teacher evaluation tool must continue to interact and complement each piece of the system.

Recommendation Three: Support for Implementation

The task force recommends ongoing support be provided within and across the system. Stakeholders emphasize communication on the purpose of the system is important. Administrators and educators should be trained together and resources should be provided for effective implementation. Stakeholders also stressed the importance of time to learn the system and implement it effectively. Ms. Darling Hammond notes strong evaluation systems need principals and other evaluators with deep knowledge of

The professional development dossier is a collection of documentation compiled to demonstrate licensure advancement competency. The collection consists of classroom data such as lesson descriptions, handouts, student work, video and audio recordings, or photos. The portfolio must meet five strands that are aligned with New Mexico teacher performance competencies. The five strands are: (1) instruction, (2) student learning, (3) professional learning, (4) superintendent verification of work, (5) recommendation based on evidence from annual teacher evaluations.

Each of the three levels of teacher licensure have nine common competencies with differentiation of performance indicators based on specific levels. The nine common competencies are:

1. The teacher accurately demonstrates knowledge of the content area and approved curriculum.
2. The teacher appropriately utilizes a variety of teaching methods and resources for each area taught.
3. The teacher communicates with and obtains feedback from students in a manner that enhances student learning and understanding.
4. The teacher comprehends the principles of student growth, development and learning, and applies them appropriately.
5. The teacher effectively utilizes student assessment techniques and procedures.
6. The teacher manages the educational setting in a manner that promotes positive student behavior, and a safe and healthy environment.
7. The teacher recognizes student diversity and creates an atmosphere conducive to the promotion of positive student involvement and self-concept.
8. The teacher demonstrates a willingness to examine and implement changes as appropriate.
9. The teacher works productively with colleagues, parents, and community members.

teaching and learning to give useful feedback to teachers and plan professional development that supports student learning. Current law requires school principals to attend a PED-approved training program to improve their evaluation, administrative, and instructional leadership skills every two years. Recently released guidance from PED requires principals to train their staff on the new system during the upcoming school year. PED has also released guidance on how principals should support their teachers in developing a professional development plan. It is unclear if PED will provide targeted training on the system this year.

Recommendation Four: Observations

The task force recommends observations be a core component of the system for all school personnel and account for 50 percent of an educator’s overall performance, including both formal and informal observations paired with immediate, actionable feedback from a trained observer. The observation process should have meaningful feedback, ensuring the focus remains on teaching and learning, not compliance. According to PED’s new guidance on the evaluation system, administrators will be expected to conduct informal observations quarterly to give quality and actionable feedback to teachers. One formal observation will also be required based on the Charlotte Danielson Framework. Statute requires the school principal to observe a teacher’s classroom practice annually to determine their ability to demonstrate state-adopted competencies. PED will need to update rule to reflect the frequency and timelines on when evaluations must be completed, including how much this element will be weighted in a teacher’s evaluation score.

In 2013, Charlotte Danielson updated her Framework to align with the Common Core State Standards.

In the previous evaluation system, teachers were rated as ineffective, minimally effective, effective, highly effective and exemplary. In the new system, teachers are rated as not demonstrating, developing, applying and innovating.

Rubric for Observations. The task force recommends evaluators use a revised version of the Danielson Framework as a rubric to observe school personnel. Revisions should include language aimed at social and emotional learning, cultural sustainability, and standards for individual education programs (IEPs). Task force members also want the rubric to include a technical guidance document to assist in observations. For non-teaching positions, the task force recommends using this rubric when appropriate and adding “look-fors” that are specific to those roles. Recently, PED released newly adapted Charlotte Danielson rubrics to rate a teacher’s performance based on four categories instead of the previous five categories.

Recommendation Five: Professional Development Plans

Current state law requires teachers and school principals to create a professional development plan for the upcoming year and the performance evaluations should be based on how well the plan was carried out. Statute requires the professional development plan for teachers include evidence that the professional development offered or required from the state is incorporated in the classroom.

Another core component of the new evaluation system is the professional development plan. The task force recommends this element should be weighted at 50 percent of an educator’s overall performance assessment. The task force recommends the professional development plan connect with other elements of the evaluation such as formal observations, student performance evidence or data, student survey data, and informal feedback from peers. Following task force recommendations, PED recently released a formal document that must be used to develop the professional development plan. As a part of the process, teachers will be asked to write a goal that is student-centered, measurable, attainable, and time-bound. The objective of the goal must be to improve teaching and student learning. Educators must explain the new skill, knowledge,



or ability they will need to meet the goal and why they chose to focus on this specific area. Possible data sources to assist in making this determination can come from self-assessment, past observations, past student data, past survey data, or a formative assessment. As a part of this process, educators must identify a timeline and specific steps they will take to reach the goal. Educators will be expected to reflect on their progress mid-year and explain at the end of the year how their progress connects their professional growth to action steps, classroom-data, and artifacts. The form does not require teachers to align their professional development plan with the nine teaching performance competencies currently included in rule, whereas the last process ensured these elements were tightly integrated into the professional development plan. PED should update rule to reflect current expectations and clarify if the professional development plan should be based on the nine teacher competencies, currently required by rule.

Evidence of Student Growth or Student Learning. Measuring student growth in relation to a teacher’s performance proved to be complicated and controversial within teacher evaluation systems across the nation. Measuring this element is necessary because student learning is the primary goal of teaching. According to Ms. Darling Hammonds research, strategies to use multiple sources of evidence about student learning are essential to get a fair gauge on what a teacher accomplished with their students.

Following federal incentives to incorporate student achievement data into state teacher evaluation systems, many states required teachers’ evaluation systems to include student test scores. One tool developed to measure student achievement was a value-added model, which compared the growth of students test scores with those of demographically similar students in other classrooms. However, many researchers found problems in the model highlighting that many factors influence student learning beyond a teacher. Studies have found up to 10 percent of student achievement can be attributed to the teacher, and 60 percent of the influences come from socioeconomic factors of students and the collective composition of the classroom or school. As of 2018, only 15 of 50 states were still using value-added models in their evaluation.

Ms. Darling Hammond asserts judgement on teachers’ contributions to student learning should rely on multiple measures of student learning, not a single test or value-added score. The use of student learning measures should take into account factors that affect student achievement gains, including student characteristics and the context of the services being offered. Evidence of student learning can be integrated into teacher evaluation systems through student learning objectives which align with PED’s current student-learning-focused professional development plan requirement. Student learning objectives, used by the New York State and Rhode Island departments of education, are goals created by teachers using data about students and their learning over a defined period of time. Options used to measure student growth through this type of strategy include rubric scored papers or projects, pre- and post-tests for units of study, end of course exams, student performance demonstration in electives, diagnostic assessments throughout the year or attainment of individualized education program goals.

Professional development is essential to educator growth and continuous improvement. Research has shown that high-quality professional development programs averaging 50 hours over a six to 12-month period increased student achievement by 21 percentage points. In comparison, less than 14 hours per year on a given topic had no effect on student learning.

Research has shown that teachers classified as “effective” using a value-added model had between a 29 percent and 59 percent chance of being classified “ineffective” the following year. An analysis of New Mexico teacher evaluation scores by Educational Policy professor and value-added modeling researcher, Audrey Amrein-Beardsley, between 2013 and 2016 showed that 28 percent of teachers’ scores varied by two or more rankings from year to year.

Recommendation Six: Surveys

The task force recommends surveys be used to inform self-reflection and professional planning, though they recommended surveys not be scored and weighted as a measure in an educator's evaluation score. The task force also recommends students should be surveyed on classroom climate and culture and families should be surveyed on school climate. In PED's guidance for the new system, surveys are one of three ways educators will receive feedback. PED highlights surveys will be a major component of the system, but the new guidance does not specify if a survey will contribute to an overall evaluation score like in previous years. With educators' professional development plans due 40 days after the beginning of the school year, it is unclear if surveys will be conducted this year. It is unclear what survey tool PED plans to use and how it will be administered. According to Ms. Darling Hammond's *Getting Teacher Evaluation Right*, student feedback about teachers' practice through surveys can lead to learning gains. For example, research on the Harvard-developed Tripod perception survey, which was a tool reviewed by the task force, was positively associated with student learning gains.

Recommendation Seven: Mentorships and Peer Observations

The task force recommends mentorships and peer observations be expanded and extended within the new teacher evaluation system. Task force members support allowing a peer observation for two of the three observations. The recommendations also stress mentorship and peer observations must support the professional development plan process. Task force members recommended expanding the mentorship program from one year to three years, with the third year providing support for licensure advancement. For educators who are new to the field or to a school district, task force members recommended a peer support program. Statute currently requires all beginning teachers to participate in a formal mentorship program, and evidence of completion of the mentorship program is a condition of licensure advancement. PED's guidance on the new evaluation system does not specify how the mentorship program will interact with the system or mention how peers will contribute to the process. Best practice around high-quality professional development programs includes mentoring where educators can collaborate with peers and learn from observing effective teachers.