New Mexico Educator Evaluation System

Recommendations from the New Mexico Teacher Evaluation Task Force

February 2020

About Region 13 Comprehensive Center and WestEd

The Region 13 Comprehensive Center is one of a national network of federally funded technical assistance centers that provide support and assistance to states in the implementation and scale-up of state education initiatives. WestEd is a nonpartisan, nonprofit, research, development, and service agency that works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults.

Michelle Lujan Grisham
Governor

Ryan Stewart, Ed.D.
Secretary of Education

Gwendolyn Perea Warniment, Ph.D.
Deputy Secretary for Teaching, Learning & Assessment
New Mexico Public Education Department

Seana Flanagan
Director, Educator Quality Division
New Mexico Public Education Department

Katya Danielle Gothie
Director, Educator Growth and Development Bureau
New Mexico Public Education Department

Task Force Members

Alyssa Agranat
Principal, Albuquerque Public Schools

Cynda Bellamy
Special Education Teacher, Albuquerque Public Schools

D'Anna Benavidez-Sullivan
Teacher, Las Cruces Public School District

Ellen Bernstein
Albuquerque Teachers Federation

Charles Bowyer
National Education Association NM

Darryl Bryant
Teacher, Albuquerque Public Schools

Lisa Burciaga Segura
Regional Education Cooperative IV

Robyn Cook
Regional Education Cooperative IX

Cari Davis
Teacher, Rio Rancho Public Schools

Jolene Delgado
Principal, Reserve Independent Schools

Kirk DeSoto
Principal, Roswell Independent Schools

Stacy Diller
Superintendent, Clayton Municipal School District

Alexis Domme
Director of Teacher Educator Programs, San Juan College

Anya Dozier Enos
Santa Fe Indian School, BIE

Candace Flint
Principal, Santa Fe Public Schools

Patrizia Flores
Teacher, Las Vegas City Schools

Task Force Members, continued

Vicky Fuessel
Regional Education Cooperative V

Lindsey Fullop
Teacher, Los Alamos Public Schools

Kelli Furney
Teacher, Gallup-McKinley Public Schools

William Hawkins
Assistant Superintendent, Hobbs Municipal Schools

Billie Helean
Teacher, Rio Rancho Public Schools

Mitchell Herz
Community, Apex Evaluation

Patricia Jimenez-Latham
Center for the Education & Study of Diverse Populations, New Mexico Highlands University

Nancy Jobe
Teacher, Dulce Independent School District

Nancy Johnson
Teacher, Carlsbad Municipal Schools

Terri Lindstrom
Teacher, Chama Valley Schools

Stephanie Ly
American Federation of Teachers

Silvia Miranda
Teacher, Clovis Municipal Schools

James Olivas
District Leadership, Bloomfield School District

Danny Parker
Assistant Superintendent, Artesia Public Schools

Mary Parr
National Education Association NM

Costa Pavlakos
Special Education, Albuquerque Public Schools

Julie Radoslovich
Principal, South Valley Academy

Faith Rosetta
Principal, Santa Fe Indian School, BIE

Ignacio Sanchez
National Education Association NM

Patricia Sandoval
Superintendent, Laguna Department of Education, Tribal BIE

Cindy Sims
Superintendent, Estancia Municipal Schools

Kathleen Swope
Teacher, Central Consolidated Schools

Edward Tabet-Cubero
Advocate, Learning Alliance

Andrea Thomas
Teacher, Central Consolidated Schools

Annjenette Torres
Parent, Santa Fe Public Schools

Gary Tripp
Cooperative Educational Services

Jeff Tuttle
Teacher, Albuquerque Public Schools

Simone Vann
Teacher, Rio Rancho Public Schools

Kathleen Wagner
Assistant Dean of the College of Education and Technology, Eastern New Mexico University

Hannah Wecks
Parent, Grant County

Contents

Executive Summary

In response to Governor Lujan Grisham's Executive Order to "work with stakeholders to identify and implement a more effective method for assessing teacher performance" (NM EO 2019-002), the New Mexico Public Education Department (NMPED) convened 13 statewide community input sessions in May and June 2019. The convenings gathered public input for re-envisioning and redesigning New Mexico's teacher evaluation system. In addition, the NMPED surveyed educators at the Inspire Teacher Conference on May 30–31, 2019. In July, the NMPED established the Teacher Evaluation Task Force with 46 members who were selected through a process that considered demographics, areas of expertise, and professional roles in order to ensure the Task Force represented a diverse array of perspectives. The Task Force met between August 2019 and January 2020 with assistance from three nonprofit, nonpartisan educational organizations that served as external partners: Learning Policy Institute (LPI), Collaborating for Outstanding Readiness in Education (CORE) at New Mexico State University, and the Region 13 Comprehensive Center at WestEd. This report summarizes the recommendations of the Task Force for the design of a new educator evaluation system that prioritizes educator professionalism and growth.

Task Force Recommendations

The Task Force created purpose statements that serve as guiding principles for the evaluation system design recommendations. Task Force members ensured that broad stakeholder input was reflected in these guiding principles, in the system design, and in the recommendations for each component of the system *and* ensured that the design and components were consistent with the stated system purpose.

Recommendation 1: Purpose of the System

The Task Force recommends that the purpose of the New Mexico Educator Evaluation System is to:

- Improve student learning, growth, and well-being;
- Promote educator learning, growth, and well-being;

- Support meaningful, actionable feedback and professional self-reflection; and
- Strengthen a learning culture through communication, collaboration, continuous improvement, and shared ownership.

The Task Force further recommends that the redesigned system be referred to as the New Mexico Educator Evaluation System to be inclusive of all educators, not just teachers.

Recommendation 2: System Design

The Task Force recommends that the system contain the following components that need to work in an integrated, coherent design:

1. *Observation and Feedback,* to be weighted at 50 percent of an educator's overall performance assessment, consisting of formal observations and supported by informal observations, both conducted by skilled and trained observers who provide immediate, actionable feedback.

2. *Professional Development Plan* (PDP), to be weighted at 50 percent of an educator's overall performance assessment, using multiple measures that include formal observations; student performance data/evidence of student learning; student survey data; and informal peer feedback to inform educator self-reflection, goal setting, and professional learning plans.

3. *Student Surveys and Performance Data/Evidence of Student Learning,* to be used to inform the PDP and not as an individually weighted score in an educator's overall performance assessment.

4. *Mentorship,* to be used formatively to provide input and support that informs professional performance and growth and thus contributes to an educator's overall performance assessment.

Recommendation 3: Support for Implementation

The Task Force recommends that ongoing support be provided within and across the system to ensure meaningful, rather than compliance-driven, implementation of the new system. The support should include the following elements:

1. Communication that clearly emphasizes the purpose of the system.

2. Training that builds the capacity of all stakeholders to use the system as intended, with administrators and educators trained together to support shared understanding and ownership.

3. Resources that support effective implementation.

4. Time for educators to learn the new system and implement it effectively.

Recommendation 4: Observations

The Task Force recommends that observations be a core component of the system for all professionals, including those in non-teaching positions.

1. Use a revised Danielson Framework with language to include social and emotional learning, cultural sustainability, standards and/or Individual Education Plans (IEPs), and a technical guidance document to be used in conjunction with observation rubrics.

2. Build the observation process on meaningful feedback and conversations that move past compliance to focus on teaching and learning.

3. Use formal observations by skilled observer(s) as a weighted measure in the summative evaluation. Use informal observations by skilled observer(s) or peers formatively for more frequent, actionable feedback, not part of the weighted observation measure in the summative evaluation.

4. For other professionals in non-teaching positions, use the Danielson Framework where appropriate and add "look-fors" (actions or practices for observers to look for) that are specific to the role or position, with local flexibility regarding who conducts those observations. Mandatory training and guidance provided by the NMPED must include how to effectively evaluate these professionals.

Recommendation 5: Professional Development Plans

The Task Force recommends that professional development plans (PDPs) be a core component of the evaluation system in which the educator owns the process. The educator is responsible for providing supporting evidence in meeting their goals.

1. Purposefully connect PDPs with all elements of the evaluation system (observations, mentorships, student performance data/evidence of student learning, and surveys), rather than having PDPs be a separate process.

2. Support educator ownership of the PDP to improve individual practice.

3. Promote the use of the PDP within a culture of collaboration that allows educators to work together.

4. Use a rubric with supporting evidence to determine the degree to which educator goals are met.

Recommendation 6: Surveys

The Task Force recommends that surveys be used to inform self-reflection and professional planning.

1. Provide a new customizable, formative, valid, and reliable survey system.

2. Use surveys to inform educator practice and development, not as a scored and weighted measure in a summative evaluation.

3. Minimize classroom time to administer the surveys.

4. Survey students about classroom climate and culture; survey families about school climate.

Recommendation 7: Mentorships and Peer Observations

The Task Force recommends that mentorships and peer observations be expanded and extended.

1. Add the option of peer observation (entire class period) for two of the three required informal observations.

2. Ensure that mentorships and peer observations support the PDP process.

3. Extend the mentorship program from one to three years, with the third year providing support for licensure advancement, contingent upon available resources. Consider allowing any educator to opt in.

4. Expand support to include a one-year advisory or peer support program for educators who are new to the district but not to the field.

Introduction

In January 2019, New Mexico Governor Michelle Lujan Grisham signed an Executive Order to "work with stakeholders to identify and implement a more effective method for assessing teacher performance" (NM EO 2019-002). The Executive Order further charged the New Mexico Public Education Department (NMPED) with pursuing ratings and assessments that offer support for educators to improve their practice, provide more time for instruction, and conserve resources. It specified a collaborative approach that would include input from the NMPED, teachers, administrators, parents, students, and recognized experts in the field of teacher assessment.

Engaging Stakeholders

In May and June 2019, the Educator Growth and Development Team at the NMPED held 13 stakeholder community input sessions across the state as part of the effort to re-envision the educator evaluation system (see Appendix A: Stakeholder Outreach Schedule). These sessions were open to educators, administrators, human resource teams, families, and community members to provide input on what components should be included in the new system, the evaluation process, and how the new system could be used to improve teaching and learning outcomes. In addition, a survey was administered to attendees at the *Inspire Teacher Conference* on May 30–31, 2019 (see session notes and survey data, available at http://www.core-nmsu.org/). These efforts were intended to engage as many stakeholder voices as possible in the design of the new system.

The NMPED contracted with the Learning Policy Institute (LPI) and Collaborating for Outstanding Readiness in Education (CORE) at New Mexico State University to assist in the facilitation and summary of findings of all stakeholder sessions. LPI summarized key themes across discussions, areas of concern, and possibilities for improvement in a July 31, 2019, memorandum (see Appendix B: Learning Policy Institute Memorandum). The dominant theme from stakeholders was the need for an educator evaluation system that prioritizes educator professionalism and growth. Key takeaways reported by LPI include:

1. Designing a fair and objective evaluation system that encourages growth, support, and collaboration.

2. Embracing actionable and meaningful measures of effective teacher practice and professionalism.

3. Assessing educators' contributions to student learning in a way that relies on multiple measures, not a single test or value-added score.

4. Differentiating evaluation by educator role and experience.

5. Building a fair and well-supported system.

The New Mexico Teacher Evaluation Task Force: Structure and Process

During the same time frame, the NMPED began to establish a Teacher Evaluation Task Force. Forty-six members were appointed by the Secretary of Education, based on several considerations that included demographics, area of expertise, and professional role in order to ensure the Task Force represented a diverse array of perspectives from across the state.

The Task Force was charged with two tasks:

1. Propose an interim teacher evaluation process for the teacher summative reports for the 2019/20 school year.

2. Provide recommendations to the Governor and the Secretary of Education for a redesigned teacher evaluation system to be implemented in 2020/21 and beyond.

The Task Force met monthly from August 2019 to January 2020, facilitated by NMPED partners CORE, LPI, and Region 13 Comprehensive Center, a federally funded technical assistance center at WestEd (see Appendix C: Task Force Schedule).

Interim System Design

The interim plan did not include student assessment data as part of the evaluation framework. It consisted of the following components:

1. Classroom Observations (Domains 2 and 3 in the Danielson Framework Observation Rubric), weighted at 50 percent

2. Planning, Preparation, and Professionalism (Domains 1 and 4 in the Danielson Framework Observation Rubric), weighted at 40 percent

3. Family or Student Surveys, weighted at 10 percent

In addition, the state's teacher evaluation rubric was revised to reduce five levels of performance to four, renaming them to: Innovating, Applying, Developing, and Not Demonstrating.

A minimum of one formal observation during the year was required along with three walk-throughs (conducted by an administrator) focused on the elements and domains specified in the rubric. Scores from walkthroughs were not used in teachers' summative evaluations.

New System Design

The Task Force began its work on the recommendations for the new system design by reviewing the Governor's Executive Order and the stakeholder input from face-to-face sessions and educator surveys; defining the purpose and intended uses of an evaluation system, which would drive its design; identifying parameters such as existing federal and state statutory requirements; and considering implementation constraints. They proceeded to define the core components of an educator evaluation system and formed working groups to more deeply examine each component and its place in the overall system. The working groups collaborated with each other and presented their work to the full Task Force for discussion and consensus-building. After several months of work, the Task Force was prepared to formulate its consensus recommendations at its January 2020 meeting.

The contents of this report are based on the considerations and consensus agreements among Task Force members. Where whole-group consensus was not reached, decisions were made based on majority approval. A representative from each working group was nominated to review and confirm the recommendations of the Task Force stated in this report to ensure fidelity to the Task Force's intentions. The following section presents the recommendations and includes a rationale for each recommendation as well as an explanation of how the Task Force envisions its implementation.

Task Force Recommendations

Recommendation 1: Purpose of the System

The Task Force recommends that the purpose of the New Mexico Educator Evaluation System is to:

- Improve student learning, growth, and well-being;
- Promote educator learning, growth, and well-being;
- Support meaningful, actionable feedback and professional self-reflection; and
- Strengthen a learning culture through communication, collaboration, continuous improvement, and shared ownership.

In reviewing the Governor's Executive Order and stakeholder input, several themes emerged regarding the purpose of the educator evaluation system. The Executive Order clearly set

forth the expectation for a system that would "offer support for educators to improve their practice"; likewise, stakeholders desired a system that "encourages growth, support, and collaboration" (see Appendix B, Key Takeaway #1). The Task Force members reflected on input from stakeholders emphasizing the need for "actionable and meaningful measures of effective teacher practice and professionalism" (Key Takeaway #2). They then reached consensus from draft vision statements that would serve as guiding principles in system design.

Task Force members affirmed a theory of action that would drive the design based on the system's ultimate goal — to improve student learning, growth, and well-being:

> *If we promote educator growth, support, and well-being through meaningful, actionable feedback and professional self-reflection in a learning culture characterized by communication, collaboration, continuous improvement, and shared ownership...*
>
> *Then student learning, growth, and well-being will improve.*

The Task Force further recommends that the redesigned system be referred to as the New Mexico Educator Evaluation System to be inclusive of all educators, not just teachers.

Recommendation 2: System Design

The Task Force recommends that the system contain the following components that need to work in an integrated, coherent design.

1. *Observation and Feedback,* to be weighted at 50 percent of an educator's overall performance assessment, consisting of formal observations and supported by informal observations, both conducted by skilled and trained observers who provide immediate, actionable feedback.

2. *Professional Development Plan* (PDP), to be weighted at 50 percent of an educator's overall performance assessment, using multiple measures that include formal observations; student performance data/evidence of student learning; student survey data; and informal peer feedback to inform educator self-reflection, goal setting, and professional learning plans.

3. *Student Surveys and Performance Data/Evidence of Student Learning,* to be used to inform the PDP and not as an individually weighted score in an educator's overall performance assessment.

4. *Mentorship,* to be used formatively to provide input and support that informs professional performance and growth and thus contributes to an educator's overall performance assessment.

Three of the four components of the recommended system are supported by New Mexico statute: Observations (NMSA 22-10A-19A&C), Professional Development Plans (NMSA-22-10A-19D), and Mentorship (NMSA22-10A-9). In accordance with statutory requirements,

the Task Force recommends these components remain the core features of the system. Task Force members were clear that an educator's evaluation should be based on multiple measures, and their recommendations include details on how those measures should contribute to an overall holistic assessment. Fifty percent of an educator's evaluation should be based on the quality and identified outcomes of a PDP, driven by educators' self-assessment and reflection on their own professional practice and continuous improvement goals. Fifty percent should be based on observation and meaningful feedback from skilled observers to provide an objective (i.e., external) assessment of professional practice and continuous improvement goals. Both of those components should be informed by informal observations, student surveys, and evidence of student learning which includes student work and assessment data.

Task Force members agreed that if these two components are equally weighted, the purpose of both, and the system itself, remains focused on professional learning and growth with the intention of improving student outcomes. Mentoring and peer observation remain as core features of the system in order to inform and support educators in meeting their professional learning goals.

Recommendation 3: Support for Implementation

The Task Force recommends that ongoing support be provided within and across the system to ensure meaningful, rather than compliance-driven, implementation of the new system. The support should include the following elements:

1. Communication that clearly emphasizes the purpose of the system.
2. Training that builds the capacity of all stakeholders to use the system as intended with administrators and educators trained together to support shared understanding and ownership.
3. Resources that support effective implementation.
4. Time for educators to learn the new system and implement it effectively.

Task Force members recognize that any design, no matter how well-intentioned, will ultimately depend on the quality and integrity of its implementation. Measuring educator effectiveness is a complex undertaking that involves the ongoing commitment of all participants. Redesigning the system provides the opportunity to re-envision what New Mexico educators value but will require thoughtful attention to the system's implementation to actualize its purpose. Transitioning to a new system will require clear communication that creates a vision for meaningful, continuous professional growth and improvement intended to improve outcomes for students. It will require accessible and ongoing training that builds local capacity, collective ownership, and transparency to implement the system effectively. In-person trainings and webinars could be held regionally throughout the year and offered on an ongoing basis to review and relearn. Doing so will require the identification and commitment of adequate resources to support implementation. And finally, the Task Force members recommend

that adequate time be given to all system participants to implement the system well and that adequate time be given to all educators to reflect upon and strengthen their practice.

Recommendation 4: Observations

The Task Force recommends that observations be a core component of the system for all professionals, including those in non-teaching positions.

1. Use a revised Danielson Framework with language to include social and emotional learning, cultural sustainability, standards and/or Individual Education Plans (IEPs), and a technical guidance document to be used in conjunction with observation rubrics.

2. Build the observation process on meaningful feedback and conversations that move past compliance to focus on teaching and learning.

3. Use formal observations by skilled observer(s) as a weighted measure in the summative evaluation. Use informal observations by skilled observer(s) or peers formatively for more frequent, actionable feedback, not part of the weighted observation measure in the summative evaluation.

4. For other professionals in non-teaching positions, use the Danielson Framework where appropriate and add "look-fors" (actions or practices for observers to look for) that are specific to the role or position, with local flexibility regarding who conducts those observations. Mandatory training and guidance provided by the NMPED must include how to effectively evaluate these professionals.

The Observation Tool Working Group provided input on the redesign of the observation and feedback process that included the use of a revised Danielson Framework; the use of formal and informal observations; and the need for meaningful feedback. Observation and feedback have been shown to be effective in improving educator practice (Taylor & Tyler 2012). Formal observations should be conducted using the revised Danielson Framework. These revisions included maintaining the renamed four levels used in the interim, and replacing subjective language with observable, clarifying language, including language about standards, IEP goals, and linguistically and culturally responsive instruction. The revised Danielson Framework should serve as the instrument for formal observations and should guide the feedback discussion.

The Observation Tool Working Group also suggested a change in terminology from "walk-throughs" to "informal observations." The purpose of informal observations is to provide formative feedback. Informal observations should identify evidence of the PDP goals, as well as overall practice; they do not receive ratings on the rubric that are averaged into the summative evaluation rating. This change is intended to address the feedback from educators that not all elements can be observed in a short, informal observation, and therefore not all elements should be scored. Narrative content from informal observations may be included in the summative performance assessment as additional evidence.

The "Other Professionals" (those in non-teaching positions) Working Group affirmed that staff in non-teaching positions also need the opportunity for self-reflection, observation, and feedback to support their professional growth; and that the qualifications for licensure should not be confused with evaluation. Competencies are what one needs to qualify for a professional license; evaluations are about performance. The working group also recognized that certain modifications may be needed because of the varied roles of non-teaching staff. The working group suggested using either an adapted Danielson Framework or a newly created one, and they spent some time developing a possible draft. Guidelines from the NMPED regarding who should conduct these observations, appropriate look-fors, degree of local flexibility, and so on, are needed for implementation decision-making at the local level.

The Task Force agreed on these recommendations and that the system needs to be built around meaningful feedback that builds trust, collaboration, and growth, with a focus on teaching and learning.

Recommendation 5: Professional Development Plans

The Task Force recommends that professional development plans (PDPs) be a core component of the evaluation system in which the educator owns the process. The educator is responsible for providing supporting evidence in meeting their goals.

1. Purposefully connect PDPs with all elements of the evaluation system (observations, mentorships, student performance data/evidence of student learning, and surveys), rather than having PDPs be a separate process.

2. Support educator ownership of the PDP to improve individual practice.

3. Promote the use of the PDP within a culture of collaboration that allows educators to work together.

4. Use a rubric with supporting evidence to determine the degree to which educator goals are met.

The Professional Development Plan Working Group recognized that in order for the PDP process to be meaningful, it must be about professional growth and learning and be focused on self-assessment and reflection. Self-assessment and reflection can serve as powerful techniques for professional growth and change in professional practice (Ross & Bruce 2007; van Diggelen, den Brok, & Beijaard 2013). There must be relevant and aligned connections between the PDP, mentorship, observations, feedback, and evidence of student learning. The PDP should not lie outside the evaluation process but be an integral part of it — all the work of the PDP should be grounded in the Danielson Framework domains and competencies. Collaboration and feedback should be continuous and timely between the educator, mentors/peers, and the evaluator in order to directly support an effective learning environment. The PDP should reflect individual as well as collaborative work evidenced by a portfolio that includes artifacts from assessment data, student work, and an application of feedback from

observations and mentors or peers. And finally, the PDP process should be guided by PDP rubric scoring that determines the degree of educator professional learning and growth and ensures that the process is implemented effectively and with fidelity.

The Task Force envisioned the PDP as a dynamic process that encourages action research, stretching one's practice, and taking risks in order to support innovation. Training, guidance, and attention to purpose will be needed to ensure that PDPs are integrated within the system, embedded in daily practice, focused on educator practice and student learning, and manageable so that the PDP does not become a burdensome exercise but rather an opportunity for innovative teaching.

Recommendation 6: Surveys

The Task Force recommends that surveys be used to inform self-reflection and professional planning.

1. Provide a new customizable, formative, valid, and reliable survey system.
2. Use surveys to inform educator practice and development, not as a scored and weighted measure in a summative evaluation.
3. Minimize classroom time to administer the surveys.
4. Survey students about classroom climate and culture; survey families about school climate.

A new customizable, formative, valid, and reliable survey system should be developed to allow educators to focus on growth based on student perception data and to build trust through collaboration between educators throughout the school. Given this purpose, survey results should not be a scored element in the summative educator evaluation. Surveys can provide rich data that educators should intentionally incorporate into informal and/or formal (i.e., PDP) continuous professional growth. Schools should administer surveys at least once each year and could have the option of administering them multiple times in a year to further inform educator practice and provide timely feedback in a formative process.

The Survey Working Group differentiated surveys that are used to assess *school* climate and culture from those designed to assess *classroom* climate and culture and *educator* practice as perceived by students. In focusing on the latter, the Working Group affirmed that the purpose of a student survey should be to inform educator self-reflection and therefore serve as an input to the PDP.

Finally, the survey system should minimize the time students spend taking a survey, easing the burden of survey administration on students, families, schools, districts, and other stakeholders. The survey system should allow for the purposeful sampling of students so all educators have a valid sample of student surveys, and the system should be flexible to allow ease of use for schools and districts of all sizes. The survey instrument provider should have a

long history of supplying responsive customer support for schools and districts and be able to support the NMPED in providing timely and adequate training for educators in the use of the survey system.

The Survey Working Group identified six priorities and values that guided their work in considering these recommendations on the use of surveys. Surveys should:

- Be used in a **formative** manner to enhance and grow educators' practice.
- Foster **collaboration** between educators and peers, evaluators, mentors, and others.
- Be **easy to implement** at the classroom, school, district, and state level.
- Provide a full picture of and **rich data** about classrooms and contexts.
- Build **trust** in the survey and the larger evaluation system among educators.
- Provide information to enhance educational **equity** in the state.

The working group examined four separate survey instruments that met these basic qualifications (Tripod 7Cs, Panorama, ED 5 Essentials, and NM Engagement), scoring each instrument on the six priorities laid out above. From the analysis, the group found that *Tripod 7Cs Student Survey* (tripoded.com) would best meet the needs for an educator survey as a formative indicator in the state educator evaluation system. However, if it is determined that *Tripod 7Cs Student Survey* is not feasible for statewide use, the Task Force recommends these six priorities be considered in the selection and the selected survey be used for formative purposes and not weighted separately in an educator's overall summative evaluation of performance.

The Task Force agreed that student voice is important and should be used to inform educator self-reflection and professional growth and not be included as a weighted score in an educator's overall performance assessment.

Recommendation 7: Mentorships and Peer Observations

The Task Force recommends that mentorships and peer observations be expanded and extended.

1. Add the option of peer observation (entire class period) for two of the three required informal observations.
2. Ensure that mentorships and peer observations support the PDP process.
3. Extend the mentorship program from one to three years, with the third year providing support for licensure advancement, contingent upon available resources. Consider allowing any educator to opt in.
4. Expand support to include a one-year advisory or peer support program for educators who are new to the district but not to the field.

The working group reviewed the current New Mexico statute and identified ways in which mentoring and peer support could be extended and expanded to provide stronger support to educators and address gaps and needs in the current process. They affirmed that mentors need to serve in a support role in the system.

The Task Force agreed that the mentorship program should be expanded from one to three years and should allow any educator to opt into the mentorship program. Research indicates many teachers leave the profession in the first three to five years (Gray & Taie 2015; Nguyen, Pham, Springer, & Crouch 2019) and the impact of mentorships on teacher attrition is well-supported in research (Ingersoll & Strong 2011). Expanding mentoring to three years will align and support the tiered licensure process (NMSA 22-10A-4). And finally, the Task Force recommends that peer observation and feedback be an option for up to two informal observations. Studies have shown that peer observation and feedback have a significant impact on teacher practice and student outcomes (Burgess, Rawal, & Taylor 2019).

Conclusion

This report summarizes the work of the New Mexico Teacher Evaluation Task Force for the development of a redesigned educator evaluation system. It represents the Task Force's deliberations on the issues, the considerations of their charge, state and federal statutory requirements, and stakeholder input. It includes both the Task Force's consensus recommendations and related implications for implementation. The Task Force addressed the system as a whole, as well as each of its component parts, through small-group and whole-group processes. As noted in this report, Task Force members recognize that any design, no matter how well-intentioned, will ultimately depend on the quality and integrity of its implementation and the ongoing commitment and shared ownership of all its participants.

References

Burgess, S., Rawal, S., & Taylor, E. S. (2019). *Teacher peer observation and student test scores: Evidence from a field experiment in English secondary schools* (EdWorkingPaper: 19-139). Providence, RI: Annenberg Institute at Brown University. Retrieved from http://www.edworkingpapers.com/ai19-139

Gray, L., & Taie, S. (2015). *Public school teacher attrition and mobility in the first five years: Results from the first through fifth waves of the 2007-08 Beginning Teacher Longitudinal Study* (NCES 2015-337). Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubs2015/2015337.pdf

Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research, 81*(2), 201-233. Retrieved from https://www.gse.upenn.edu/pdf/rmi/RER_article.pdf

Nguyen, T. D., Pham, L., Springer, M., & Crouch, M. (2019). *The factors of teacher attrition and retention: An updated and expanded meta-analysis of the literature* (EdWorkingPaper: 19-149). Providence, RI: Annenberg Institute at Brown University. Retrieved from https://edworkingpapers.com/ai19-149

Ross, J. A., & Bruce, C. (2007). Teacher self-assessment: A mechanism for facilitating professional growth. *Teaching and Teacher Education, 23*(2), 146-159. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0742051X06000813

Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review, 102*(7), 3628-3651. Retrieved from https://faculty.smu.edu/millimet/classes/eco7321/papers/taylor%20tyler%202012.pdf

Tripod Partners. (2016). *Tripod's 7Cs framework of effective teaching*. Retrieved from https://tripoded.com/wp-content/uploads/2017/01/Guide-to-Tripods-7Cs-Framework-of-Effective-Teaching.pdf

van Diggelen, M., den Brok, P., & Beijaard, D. (2013). Teachers' use of a self-assessment procedure: The role of criteria, standards, feedback and reflection. *Teachers and Teaching: Theory and Practice, 19*(2), 115-134. Retrieved from https://eric.ed.gov/?id=EJ1012498

Appendix A:
Stakeholder Outreach Schedule

Re-envisioning New Mexico's Educator Effectiveness System

STAKEHOLDER COMMUNITY INPUT SESSIONS

The New Mexico Public Education Department (NMPED), in partnership with New Mexico State University, invites teachers, administrators, human resource teams, families, and community members to provide input on how best to measure educator effectiveness in New Mexico. Community conversations will be held at the following locations, dates, and times.

Tuesday, May 28, 2019
BERNALILLO
Bernalillo High School Cafeteria
148 Spartan Alley
4:30 - 7:00 PM

Monday, June 3, 2019
SANTA FE
Santa Fe Public Schools
BF Young Sierra Vista Room
1300 Camino Sierra Vista
4:30 - 7:00 PM

Tuesday, June 4, 2019
ESPAÑOLA
Northern NM College
921 Paseo de Oñate
4:30 - 7:00 PM

Wednesday, June 5, 2019
FARMINGTON
FMS Conference Center
1400B E. 20th Street
4:30 - 7:00 PM

Thursday, June 6, 2019
GALLUP
Gallup Schools Board Room
640 Boardman Dr.
4:30 - 7:00 PM

Monday, June 10, 2019
DEMING
Southwest REC
1321 E. Poplar
4:30 - 7:00 PM

Tuesday, June 11, 2019
LAS CRUCES
CPS Conf. Rooms A/B
505 S. Main St, Suite 249
4:30 - 7:00 PM

Wednesday, June 12, 2019
RUIDOSO
Ruidoso Middle School
123 Warrior Dr.
4:30 - 7:00 PM

Monday, June 17, 2019
HOBBS
Hobbs Schools Training Center
2110 E. Sanger St.
4:30 - 7:00 PM

Tuesday, June 18, 2019
CLOVIS
Clovis High School Cafeteria
1900 N. Thornton St.
4:30 - 7:00 PM

Wednesday, June 19, 2019
LAS VEGAS
NMHU Student Union Building
Room 322
800 National Ave.
4:30 - 7:00 PM

Thursday, June 20 and Friday, June 21, 2019
ALBUQUERQUE
Crowne Plaza Albuquerque
1901 University Blvd. NE
4:30 - 7:00 PM (**Thursday June 20, 2019**)
2:00 – 4:30 PM (**Friday June 21, 2019**)

If special accommodations, ASL, or Spanish translation is required, please contact Region 9 at 575-257-2368

For questions, contact seana.flanagan@state.nm.us
For additional information on how your input will be used to develop a statewide teacher evaluation system, visit www.core-nmsu.org

Appendix B:
Learning Policy Institute Memorandum – "New Mexico Evaluation Outreach Themes," July 31, 2019

TO: New Mexico Public Education Department
FROM: Learning Policy Institute
RE: New Mexico Teacher Evaluation Outreach Themes
DATE: July 31, 2019

Between Tuesday, May 28, and Friday, June 21, 2019, the New Mexico Public Education Department (PED) held 13 stakeholder community input sessions across the state as part of efforts to re-envision the educator effectiveness system. These sessions were open to teachers, administrators, human resource teams, families, and community members to provide input on what components should be included in the new teacher evaluation system, the evaluation process, and how the new system should be used to improve teaching and learning outcomes. (See outreach schedule here.)

This memo presents several takeaways from these sessions: key themes across discussions, areas of concern, and possibilities for improvement. It was drawn from the session notes and posters collected by the CORE team from New Mexico State University, follow-up interviews with the CORE team and members of the PED who attended all or most of the sessions, and the observation notes compiled by members of the Learning Policy Institute team from the sessions each attended.

The dominant theme of these meetings is that New Mexico's education stakeholders want a teacher evaluation system that prioritizes teacher professionalism and growth—not punishing and embarrassing teachers. Teachers are ready to move to a system that values the full scope of work they do every day in the classroom, encourages them to take risks to improve student learning, recognizes their strengths and achievements, and gives them opportunities to improve their practice and collaborate with their peers and administrators. Administrators agree with this vision for teacher evaluation that helps them better understand the strengths of their teachers and how they can support teachers' development, but they need a more streamlined, easy-to-use system that allows them to spend more time in classrooms and building positive relationships with teachers and less time sitting in front of a computer, uploading reports. Community members and advocates also want a system that respects teachers and keeps them in the classroom, values their efforts to connect to their communities, and is clear and reasonable.

There is also broad understanding from all stakeholder groups that the transition to this new system will not be easy, but they see the benefits—a more authentic process, stronger relationships, and putting the focus back on cultivating good teacher practice to improve student learning—as far outweighing the challenges that come with change. However, it will be critical that the Task Force take the following into consideration during the transition to the new system:

1. The Task Force should acknowledge the lingering tension and distrust between teachers and administrators and the PED and make restoring trust an integral part of the process. The stakeholder engagement meetings were a good start to rebuilding trust, but the Task Force and the PED should understand that this will be an ongoing process and find ways to include teacher, administrator, and community member voice over the course of the Task Force's work and into implementation of the new system.

1

2. The Task Force might consider how to be as transparent as possible throughout this process. All decisions must be communicated in a timely, effective, and open manner that allows stakeholders to fully understand how decisions were reached, and where decisions are made that diverge from stakeholder opinion, the reasoning behind them should be fully publicized. This will go a long way toward ensuring that trust is built and maintained throughout the process, which is something many teachers felt was lacking in recent years.

3. The transition to the new system will require training for administrators and teachers on all aspects of the new system, but both groups said the most important piece will be training on the new components of the observation rubric and the process for implementing the system effectively. Trainings that bring together administrators and teachers can help build a level of trust and put both groups on the same page in terms of what is expected in the new observation process as well as what defines excellent teaching.

4. Beyond training for teachers and administrators, the transition between the old teacher evaluation system and the new one will require a mindset shift about the purpose of teacher evaluation. The PED might consider how to create an explicit and ongoing dialogue on how moving from a punitive system to a growth system creates an entirely new paradigm for what is valued and important in schools. In addition to redesigning the teacher evaluation system, the Task Force and the PED will likely want to create a communications strategy that helps all stakeholders understand this shift and clear messaging from the PED that they are fully committed to this change and supporting educators and school staff during the transition.

5. There remains a gap in teacher and administrator knowledge about what is state-controlled in the evaluation system and what is up to local determination. The Task Force and, ultimately, the PED will need to explicitly explain where state parameters are in place and how local flexibility is enabled on various components of the evaluation system. The Task Force will also need to acknowledge the tension between wanting a level of local flexibility in the evaluation process while maintaining the need for consistency in implementation and determining progression levels across the state.

6. It will be useful for the PED and its Task Force to remind stakeholders to maintain realistic expectations of the time and capacity-building needed for successful changes and understand that change will not happen overnight. Teachers and administrators are already asked to do so much, and to gain buy-in, a plan must be made to systematize the new teacher evaluation process in a way that does not feel like "one more thing."

KEY TAKEAWAYS

Below we have grouped stakeholder comments regarding a re-envisioned teacher evaluation system by five major takeaways about what stakeholders hope for:

Key Takeaway 1: Designing a Fair and Objective Evaluation System That Encourages Growth, Support, and Collaboration—Not Punishment

The top priority of all stakeholders at every meeting across the state is to create a teacher evaluation system that prioritizes teacher growth, development, and professionalism and

2

contributes toward an environment of collaboration and support. Teachers, administrators, and community stakeholders want a system that acknowledges the full range of what teachers do every day in their classrooms as well as to create a positive learning environment across the school. Additionally, they want a system that serves as a meaningful and authentic tool to recognize teacher strengths, identify areas of needed support, and create professional development plans aligned with teacher and school needs.

Key Takeaway 2: Embracing Actionable and Meaningful Measures of Effective Teacher Practice and Professionalism

To create this positive, growth-focused teacher evaluation system, stakeholders agree consideration must be given to all of the following elements:

- alignment of evaluations to professional development and support
- clear and specific observation rubrics focused on important dimensions of in-classroom practice
- multiple, ongoing observations to document growth and provide useful, timely feedback
- opportunities for teacher self-reflection
- inclusion of formative teacher data (e.g., lesson or unit plans, professional development, artifacts of student work)
- student and parent surveys used for improving individual practice and school climate; and
- using teacher attendance as an incentive, not for punishment.

Key Takeaway 3: Assessing Teachers' Contributions to Student Learning Should Rely on Multiple Measures, Not a Single Test or Value-Added Score

Stakeholders also believe that a teacher evaluation system built for growth and support must include:

- valid student assessment data that provides a measure of student learning growth over a single year and allows teachers to document their students' learning and
- authentic evidence of student learning.

Key Takeaway 4: Differentiating Evaluation by Teacher Role and Experience

Stakeholders feel that a teacher evaluation system useful for teachers' individual and collective growth should be differentiated in the following ways:

- by level of experience,
- by students served, and
- by content area.

Additionally, stakeholders articulate an interest in having local flexibility within the new teacher evaluation system in order to determine and support local teaching priorities.

3

Key Takeaway 5: Building a Fair and Well-Supported System

In addition to the above elements, stakeholders feel that several additional elements would be necessary for the creation and implementation of an effective, growth-oriented teacher evaluation system. These include:

- widely available training to support transition to the new system
- a manageable and user-friendly evaluation process
- access to resources and exemplars that demonstrate levels of teaching defined in rubrics, and
- a fair appeals process to address disagreements and/or unreflective observations.

Within this analysis, we have included an overview for each takeaway that further explains stakeholders' views on the theme, as well as stakeholder quotes to highlight specific, illustrative points of feedback. We attribute quotes to whole groups, rather than individuals, because the comments represent consensus viewpoints collaboratively developed in role-alike groups and shared during the group share-out portion of stakeholder meetings.

Key Takeaway 1: Designing a Fair and Objective Evaluation System That Encourages Growth, Support, and Collaboration—Not Punishment

"Get rid of teacher evaluation that pushes teachers out of the profession."
- *District Administrators, Farmington Stakeholder Meeting*

"Celebrate what teachers are good at—[we] need to know where our strengths are and maximize them, as well as find the areas where growth is needed."
- *Secondary Educators, Ruidoso Stakeholder Meeting*

There is overwhelming agreement among all groups of stakeholders and in all regions of the state that a new teacher evaluation system should be focused on helping teachers to grow and improve their practice—not on punishing teachers or pushing teachers out of the profession. Teachers, administrators, and community members agree that the old system focused far too much on criticizing teachers without any opportunities for improvement or direct pathways to support.

Teachers are highly receptive to an evaluation system that offers continuous, constructive feedback and coaching; helps them to understand their strengths and areas for growth; and provides pathways to leadership roles in their school. Similarly, administrators seek a system that presents a more authentic process that captures effective teaching and helps them understand how to better retain and support teachers. Teachers specifically asked for an evaluation system that celebrates and maximizes their strengths, recognizes their growth, and encourages them to take chances to improve their students' learning.

4

One specific request to reduce the stigma from the former system is to replace existing rubric levels (currently designated as Ineffective, Minimally Effective, Effective, Highly Effective, and Exemplary) to another set of indicators showing progression (e.g., numbered levels 1–4 or descriptors such as "novice" through "expert"). Stakeholders noted that educators often associate the 5-tier rating system with an A–F system. From this perspective, an "Effective" rating corresponds with a C:

"Teachers use a 5-point scale to measure themselves and compare 'Effective' to a C. To avoid this, consider using a 4-point system."
- *District Administrators, Deming Stakeholder Meeting*

Other stakeholders recommended using a dashboard instead of a summative report to emphasize support and growth and give teachers and administrators a clear picture of all of the elements of the evaluation and the ability to easily pinpoint strengths and weaknesses. Teachers are also strongly in favor of replacing the current static evaluation with a "living document" that they can add to throughout the year, use to monitor their progress toward their goals, and build on from year to year.

"Administrators should have an authentic understanding of the professionals in their building and their teachers' strengths and areas for growth."
- *Elementary Educators, Las Cruces Stakeholder Meeting*

"What is the outcome? What can you do to help me grow?"
- *Secondary Educators, Farmington Stakeholder Meeting*

"Evaluation should be an opportunity to support growth and recognize teacher achievement."
- *Secondary Educators, Clovis Stakeholder Meeting*

There is also overwhelming agreement on the need for transparency and objectivity in the new teacher evaluation system to ensure that all stakeholders have complete clarity on what teachers are being evaluated on and to reduce the impact of bias on evaluations. There is concern among teachers that having a strained relationship with their administrator may skew scores negatively on their evaluation, particularly during observations, and that there should be measures in place, including bringing neutral observers into the process and triangulating observation outcomes and other determined input data, to combat any bias that may occur. Additionally, there are calls from all groups of stakeholders to ensure the components and expectations included in the evaluation system are clear to everyone, especially teachers, so that there are no surprises during the process and, ultimately, when teachers receive their final evaluation results.

"[We want] clear communication about elements of teacher evaluation…. What is expected?"
- *Elementary Educators, Santa Fe Stakeholder Meeting*

5

Both teachers and administrators are also eager to have an evaluation system that promotes greater levels of collaboration between teachers, between teachers and administrators, and between schools and the community. Teachers and administrators expressed frustration with the current system, which is perceived as fostering a competitive and divisive environment, and they, along with community members, want a system that encourages more positive teacher–administrator relationships, greater levels of cooperation between teachers, and increased opportunities for all adults in and outside of the school to work together to improve teaching and learning for all students. Teachers are particularly keen on having opportunities to conduct learning observations in other teachers' classrooms to gain new skills and learn new teaching methods, as well as more time to work in vertical and horizontal teams.

"[We want] collaboration instead of competition among teachers."
- *Elementary Educators, Albuquerque Stakeholder Meeting*

"[Evaluations should] promote collaboration … collaboration between teachers, between teachers and administrators, between teachers and instructional coaches."
- *Administrators, Las Vegas Stakeholder Meeting*

"[We] think that teachers should be able to support each other and be able to help coach each other."
- *Community Members, Hobbs Stakeholder Meeting*

Key Takeaway 2: Embracing Actionable and Meaningful Measures of Effective Teacher Practice and Professionalism

In designing a new teacher evaluation system, stakeholders offered several suggestions on the Danielson Framework and measures of teacher practice and professionalism—observations, self-reflections, artifacts, surveys, and attendance—that they believed would be fair, objective, and focused on supporting teachers' growth and development. Among the areas of greatest stakeholder consensus are the need to use observations to document growth and provide useful, timely feedback; allow for multiple perspectives in observations, including peers, external observers, and students; make space for teacher self-reflection in observations; include artifacts that are representative of authentic teacher classroom practice; and find more appropriate ways to use student and parent surveys and teacher attendance in the evaluation system.

Alignment of Evaluations to Professional Development and Support

Above all, stakeholders collectively agree that all components of teacher evaluations, from observations to surveys to evidence of student learning and effective teacher practice, should be aligned with professional development and support. Many are supportive of enhancing and using professional development plans (PDPs) to help teachers identify their growth goals and

6

measuring progress toward these goals in evaluations. Administrators and teachers envision a system that succinctly ties areas identified for growth, especially for novice teachers, to a plan of action for support and improvement. Most feel that the current system does not help administrators understand the needs of their teachers, which leads to a lack of support for teachers and unfocused, inconsistent schoolwide professional development that does little to advance teacher development or school improvement goals. Teachers are also frustrated by what they see as stand-alone observations and overall evaluations that are disconnected from their professional development plans.

"Evaluation should be followed by effective professional development based on teacher need as evidenced by evaluation."
- *Board Member, Ruidoso Stakeholder Meeting*

"[We want evaluation] that is content/program specific … [and] want to avoid district-mandated professional development and have professional development that is working on individual professional development that is specific to the needs of teachers."
- *Secondary Educators, Las Cruces Stakeholder Meeting*

Clear and Specific Observation Rubrics

"Streamline domain indicators to minimize and avoid redundancy."
- *Elementary Educators, Las Cruces Stakeholder Meeting*

There is little pushback to the Danielson Framework from teachers and administrators, especially on Domains 2: Planning and Preparation and 3: Teaching for Learning), but they do see room for streamlining some of its content and improving how it is used in observations. One suggestion with widespread support is to remove the "Exemplary" category from the observation rubric because it is often looking for teacher practice that happens outside of the classroom. For example, under the current NM TEACH observation rubric, to earn an "Exemplary" rating in Domain 1D requires the teacher showing that he or she "actively engages colleagues and provides resources to them in areas that are pertinent to their needs" and "provides and trains staff for school-wide initiatives." Most teachers agree that this part of their work is important, but they feel that it should live outside of the classroom observation rubric.

Similarly, teachers understand the importance of Domain 4: Professionalism, which measures teachers on communicating with families, participating in a professional community, keeping accurate records, reflecting on practice, and displaying professionalism, but they feel that including this in the observation rubric complicates the observation process. Instead, they suggest that these measures be included elsewhere in the evaluation, potentially as items that help building administrators consider teacher leadership roles, and they should be worked on in consultation with administrators.

As one educator group from the Española stakeholder meeting put it: "Areas to be observed should be observable in the classroom."

"Domains need to be modified and simplified."
- *Building Administrators, Ruidoso Stakeholder Meeting*

Multiple, Ongoing Observations to Document Growth and Provide Useful, Timely Feedback

"[We want an] observation model that is continuous, authentic, and simple."
- *Building Administrators, Clovis Stakeholder Meeting*

Improving observations and making the process more authentic is a primary area of concern for stakeholders, and particularly for teachers and building administrators, who want formal and informal observations and walk-throughs to happen on an ongoing basis throughout the school year. Almost all agreed that multiple observations and walk-throughs across the course of the school year would help administrators gain a more authentic picture of what teachers are doing in the classroom, instead of a one-time snapshot, as well as allow both teachers and administrators the opportunity to document growth over time.

"Administrator observations are important…[and should include] clear rubrics, multiple observations, timely feedback with clear path to improvement."
- *Elementary Educators, Farmington Stakeholder Meeting*

"Observations … need to allow for a view of the teacher over time and not merely a moment in time."
- *Elementary Educators, Albuquerque Stakeholder Meeting*

Teachers and administrators are also strongly in favor of having several observations and walk-throughs or visits throughout the school year because they see a strong need for immediate feedback and creating on ongoing dialogue around specific strengths and areas for growth. Many teachers expressed frustration over receiving feedback from observations well after they were conducted and from the evaluation as a whole at the end of the school year and being unable to act on it until the next school year. Teachers are also highly in favor of having more points of contact with their building administrators before formal observations take place (e.g., two or three informal walk-throughs and conversations) to build a positive working relationship and discuss classroom context and specific identified areas for growth that can focus the formal observation.

"We would like multiple walk-throughs throughout the year, and we want feedback from each of these visits."
- *Secondary Educators, Hobbs Stakeholder Meeting*

"We want timely feedback on our walk-throughs and observations. Please give us feedback that is timely, not just at the end of the year."
- *Elementary Educators, Hobbs Stakeholder Meeting*

8

"Feedback and reflection should be a collaborative dialogue between administrators and teachers instead of just a rubric."

- *Administrators, Las Cruces Stakeholder Meeting*

To make the process more manageable, suggestions were raised in several stakeholder meetings to streamline observations by only conducting evaluations for state accountability purposes for consistently high-performing teachers every three years. Additionally, those evaluations would focus on just one agreed-upon domain or specific strands within a domain, based on areas for growth mutually identified by teachers and administrators in professional development plans. Some suggested that training highly rated, experienced teachers to become observers would also help lift the burden on administrators, and that these trained teachers could conduct observations, assist with coaching, and establish a peer observation and/or coaching system with other consistently high-performing teachers during non-high-stakes evaluation years.

There is also interest in involving the perspectives of external observers and students in the observation process. Some teachers also recommended bringing in content-specific observers who can better understand their teaching methods or having neutral, external observers brought in to conduct unbiased observations, though there is concern that these outside observers could lack the critical context of the school and/or classroom. Some teachers also discussed building in formal or informal conversations with students during observations to help administrators understand whether their experience during the observation is in sync with what students experience on a daily basis or whether it is just a "dog and pony show."

Opportunities for Teacher Self-Reflection

Several groups of teachers recommended having the opportunity to record observations to provide another level of objectivity and support self-reflection. Teachers feel that being able to record their classroom during observations and use the video during the post-observation meeting to walk through the lesson together with their observer could serve as a powerful learning tool and an opportunity to drive a shared sense of ownership over the observation. Videotaping observations or other lessons teachers want to learn from could also provide a chance for teachers to self-reflect and evaluate themselves, and the resulting self-reflection could even be the catalyst for observation debriefs, peer learning and collaboration, and professional development around building meaningful self-reflection practice.

Inclusion of Other Formative Teacher Data (e.g., Lesson Plans, Professional Development, Classroom Artifacts, Examples of Assignments or Student Work)

"Please include lesson planning and professional development in our evaluation. We spend time on these things and we want credit for these things."

- *Elementary Educators, Hobbs Stakeholder Meeting*

9

Stakeholders also discussed the inclusion of other teacher data in the evaluation system, because, as referenced in the quote above, teachers want the ability to showcase items such as lesson plans, artifacts of teaching practice and evidence of student learning, and evidence of professional development and growth aligned with their professional development plans and needs of their school. Likewise, administrators spoke of a need for multiple measures of effective teaching practice outside of observations and student learning outcomes. Teachers do ask that the new system reduces the amount of time it takes to upload these other pieces of evidence into the online evaluation portal.

"From a building administrator standpoint, we should be able to have many measures that show how a teacher is doing."
- *Building Administrators, Hobbs Stakeholder Meeting*

Student and Parent Surveys Used in Helpful Ways

"Parent/student data is very important [but] the data should not be used as punitive."
- *Community Members, Albuquerque Stakeholder Meeting*

The prevailing notion on the current iterations of the student and parent surveys is that they should not be used for accountability purposes because these surveys became less valid and reliable when there are high stakes attached. Chief among the concerns is that many teachers feel there are too many opportunities for bias in the current surveys, and they are concerned that parents and students who may not like the teacher or who hold a grudge may take this out on them in the survey. Many teachers also expressed frustration that the parent survey can be taken multiple times, which allows for negative opinions from just a few parents to dominate their survey data.

Despite these concerns, many stakeholders suggested that student and parent surveys, with more useful questions and a greater focus on the schoolwide learning environment, could be used by individual teachers for improvement of teaching practice and for both teachers and administrators to gain a better understanding of the broader school climate. One teacher at an Albuquerque stakeholder meeting discussed a personally developed student survey she uses in her classroom as the best thing she has ever done in her career because it gives her the feedback she needs to adjust her practice and helps her better understand her students' needs. Other stakeholders responded positively to the use of a student survey in this manner and the ability to personalize the survey to their own classrooms, and there was some discussion around using personalized student surveys like this as a means to show teacher reflection and willingness to grow within an evaluation system. Developing a set of collective or common questions for a student survey that teachers could draw from could also create an opportunity for shared learning among teachers around most effective practices for students.

10

Using Teacher Attendance as an Incentive, Not for Punishment

Teachers are strongly against the use of attendance data in the evaluation system because they believe it judges them on something they may not always be able to control. For example, teachers feel it is unfair to hold someone accountable for absences when unexpected health issues occur. On the other hand, building and district administrators both see a measure of attendance as important because, as one said, "Students are getting the best when teachers are in the classroom." Both groups did agree, however, that attendance could be used as an additional measure to provide teachers with high attendance "bonus points" on their overall evaluation score. There was also a small number of stakeholders who felt attendance should fall under local control, with district administrators being able to set their own threshold for teacher absences.

Key Takeaway 3: Assessing Teachers' Contributions to Student Learning Should Rely on Multiple Measures, Not a Single Test or Value-Added Score

The primary point of agreement on using student achievement data in the teacher evaluation system is that student growth should be included, but the current use of the value-add model (VAM) that compares student learning growth from year to year is unfair and unhelpful. Most stakeholders agree that regardless of the assessment chosen, it needs to happen more than once a year so that teachers are able to document the learning growth of their students across the course of the school year on what they are expected to teach and what skills and competencies their students are expected to develop. Many teachers and administrators think that at the very least, assessments should be given at the beginning of the year (e.g., a pre-test) and end of the year, but many felt that assessments that provide more immediate feedback were far more preferable.

Though there was a lot of conversation at the stakeholder meetings about a statewide student assessment, many expressed a need for having some level of local control or input on assessments while balancing the need for statewide consistency. Many feel that the use of in-classroom pre- and post-tests as an artifact of student learning growth should be considered, particularly since they offer a more immediate opportunity for teachers to understand how students are progressing in their class. There are also significant concerns about using student data on English language arts and math assessments for all teachers, particularly for electives teachers, who believe they should be measured on student outcomes that are more closely aligned with their course objectives.

"Include areas of student achievement ... student growth over time, not just three data points or just one year's growth."

- *Elementary Educators, Farmington Stakeholder Meeting*

"We want the opportunity to show growth, not proficiency."
- *Elementary Educators, Hobbs Stakeholder Meeting*

"[We think] baselines for growth, with students' pre- and post-test, makes sense within a year."
District Administrators, Gallup Stakeholder Meeting

11

"[We want] content-specific, content-related data."
- *Secondary Educators, Clovis Stakeholder Meeting*

There is widespread agreement among teachers that there needs to be a significant place in the new evaluation system to document student learning in ways other than standardized test scores. Some stakeholders discussed the use of authentic evidence of student learning in the evaluation system and allowing teachers to produce evidence and artifacts of student growth. Many teachers suggested that there should be some way of including student progress toward achieving IEPs for special education teachers. Some teachers felt that portfolios of student work should be included.

"[There should be] opportunities to demonstrate student growth through documentation, not state testing."
- *Elementary Educators, Clovis Stakeholder Meeting*

"Student achievement should be valued, but it cannot be a single snapshot of one test to determine teacher effectiveness."
- *Secondary Educators, Clovis Stakeholder Meeting*

Many stakeholder groups are also eager to determine a way to document students' social and emotional growth and development and include it in the evaluation, though they do not advocate for testing on these skills. Some suggested this could be included in teacher's PDP goals (e.g., a teacher may choose to evaluate themselves on how they contribute to students' social and emotional growth or how they include students' social and emotional development within their teaching practice). Many stakeholders also believe the new teacher evaluation system should reflect the cultural diversity of students and be inclusive of teachers' efforts to make their classrooms culturally relevant.

"[We think the] social-emotional needs of students [should be] recognized."
- *Community/Business Stakeholders, Española Stakeholder Meeting*

Key Takeaway 4: Differentiating Evaluations by Teacher Role and Experience

"One evaluation system does not fit all populations and communities."
- *Administrators, Española Stakeholder Meeting*

"Differentiate evaluation for teachers like we expect teachers to do for students."
- *District Administrators, Farmington Stakeholder Meeting*

Educators across the state repeatedly stated that there should be some level of differentiation of evaluation for teachers. Stakeholders felt that an overly rigid, unadaptable evaluation system leads to some criteria that are irrelevant, unattainable, or unproductive to an individual teacher's practice. Frequent dimensions of differentiation noted by stakeholders were teacher experience and/or licensure level, content area, and students served.

12

Additionally, other stakeholders spoke more generally of the importance of local flexibility and decision-making around the teacher evaluation system.

Differentiation in Rubric and Process by Level of Experience

At the heart of many stakeholders' interest in creating a teacher evaluation system that takes into account a teacher's experience is creating a system that is more supportive, particularly for novice teachers.

"New teachers haven't felt supported because they aren't trained in the evaluation system. We need a system for new teachers where they understand what is expected of them. They need to understand what the domains mean. We also need to provide multiple mentors for new teachers."

- *Elementary Educators, Deming Stakeholder Meeting*

"We should be allowing for developmental stages of teaching. Novices shouldn't have the same number of categories as more experienced teachers.... We should not be grading them down due to a lack of experience."

- *Elementary Educators, Albuquerque Stakeholder Meeting*

Stakeholders were strongly in favor of a system that focuses on supporting and developing novice teachers, rather than one that demoralizes them and risks pushing them out of the profession. Recommendations include orienting new teachers on the expectations of the evaluation system and adapting the evaluation system into a coaching system for teachers in their first several years of teaching. The latter would avoid perceptions of punishing novice teachers while they develop their skills. Some groups expressed interest in evaluations not being a rigid process. Instead, the system would be differentiated for teachers based on their needs (e.g., frequency of observations would be flexible). Less frequently, some called for a grace period before teachers are held accountable for student achievement outcomes. Regardless of the exact approach taken, stakeholders stressed the importance of intentionally providing novice teachers—through the evaluation system—induction-like supports, such as access to experienced mentors, coaching, and additional attention in order to meets their needs.

Differentiation by Students Served

"Special circumstances of students should be considered (e.g., IEPs, BIPs)."
- *Elementary Educators, Española Stakeholder Meeting*

"There should be some mechanism in the evaluation system that values the kids you are working with. Classroom populations vary, and teachers are not accommodated in the teacher evaluation system for the makeup of students they teach.
- *Building Administrators, Farmington Stakeholder Meeting*

"Special education and ELL teachers should be evaluated by admin and teachers that understand these types of environments in order to take into account the special circumstances that are in these types of classrooms."

13

Multiple stakeholder groups raised the point that, for particular high-need student populations (e.g., medically fragile, visually impaired, special education, and gifted and talented education students), unique domains should be developed, communicated, and aligned statewide. They felt the needs of these student groups can be so unique and varied that a one-size-fits-all evaluation framework is not useful to assess their teachers. It would not help teachers improve their instructional practice because it would not reflect the needs and learning experiences of the students they serve.

Differentiation by Content Area

"Rubrics should be available for all subjects. Differentiate by what people are actually teaching."

- *Building Administrators, Ruidoso Stakeholder Meeting*

Some stakeholders felt that evaluations should have some degree of differentiation by content area. At the greatest degree of differentiation, stakeholders requested that rubrics be content-specific. Others recommended evaluators receive proper guidance, support, and training to understand high-quality instruction within whatever subject they are evaluating a teacher in. It was also suggested by a couple groups that differentiation by content area be addressed by allowing for peer observations and feedback, especially when an evaluator is not proficient in a given subject area.

Still, while some stakeholders recommended differentiation in evaluations through content-specific rubrics or different processes based on licensure level or experience, another sizable number of groups recommended domains and rubrics that were streamlined in order to be used with all teachers.

"Rubrics should be small enough and simple enough in scope that they can be used with all teachers. The goal should be flexibility that lends itself to differentiation for grade level and licensure level."

- *Elementary Educators, Ruidoso Stakeholder Meeting*

In this case, domain rubrics would be broad and flexible, placing the responsibility on well-trained evaluators to responsibly apply them to the classroom context regardless of content area, grade level, or teacher experience.

Allowing for Local Flexibility

"Districts should have a portion of the evaluation system that they can determine."

- *Secondary Educators, Las Vegas Stakeholder Meeting*

14

"Give some power back to teachers and schools. Give them authority over, say, 10% of the evaluation. Empower schools to focus on what they need, especially in terms of professional development."

> -   *Building Administrators, Farmington Stakeholder Meeting*

Another dimension for differentiation in evaluation that stakeholders recommended was allowing for local autonomy. Some sought autonomy in terms of measuring student assessment or in determining the appropriate approach to teacher attendance, as noted previously. Several other groups around the state requested the authority to be able to determine how a portion (e.g., 10%) of their teachers' evaluation ratings are determined. One example raised was that if a particular district is focusing on including students receiving special education services into regular education classes, then evaluations could focus extra attention on how teachers serve their students with special needs.

Consistent with earlier quotes in this key takeaway, others spoke about local flexibility in the sense of being able to adapt rubrics to their unique context, whether that be in terms of student need, program assignment, or teacher experience. Understood in this way, there was broad interest in creating a responsive teacher evaluation system that reflects the diverse teaching and learning contexts in the state.

Key Takeaway 5: Building a Fair and Well-Supported System

In addition to the prior takeaways, stakeholders shared multiple suggestions to inform the redesign of the teacher evaluation system. These suggestions address process and implementation (widely available training to support transition; the importance of a manageable and user-friendly process; inclusion of a fair appeals process to address disagreements) as well as components of the system (availability of resources that demonstrate teacher performance levels).

Widely Available Training to Support Transition to New System

"Teachers and administrators [should be] co-trained for observations."

> -   *Elementary Educators, Española Stakeholder Meeting*

"All educators should be trained on evaluation criteria, not just the administration. This would act like a system of checks and balances to help ensure transparency."

> -   *Secondary Educators, Las Cruces Stakeholder Meeting*

"[There should be] clear standards and expectations—everybody is clear on what they are going to be evaluated on and in the domain language and what it looks like."

> -   *Secondary Educators, Ruidoso Stakeholder Meeting*

In community input sessions across the state, stakeholders strongly supported both teachers and administrators receiving training on the new teacher evaluation system. One important reason for this expanded training is that, in order for it to be truly focused on growth, teachers need to

15

clearly understand what they are being evaluated on as well as how to improve. In this way, the desire for training aligns to the recurring theme that this system should be oriented toward teacher development rather than operating as a punitive "gotcha" system.

As one group of secondary teachers in Las Cruces articulated, "Teachers should be aware of 'look-fors.' Train teachers on the evaluation system as well [as administrators] so everyone knows the process." This point was further emphasized by a group of building administrators in Farmington, who said, "Teachers and administrators need to have the same vision and interpretation of the science and art of teaching. Training is necessary [for that]."

It should also be noted that several groups voiced an interest in remaining engaged in the development of the re-envisioned teacher evaluation system after the community input sessions conclude. Teacher and administrator training provide one possible avenue for this.

Manageable and User-Friendly Evaluation Process

"[Evaluations should be] manageable and effective so principals are not burnt out, teaching improves, and we grow as a community."

- *Building Administrators, Farmington Stakeholder Meeting*

"[The evaluation process] should be user-friendly. Teachers should be able to use data and apply it."

- *Community Members, Bernalillo Stakeholder Meeting*

Reducing the overall burden of the teacher evaluation system, especially on building administrators, was a consistent theme across the state. Many groups expressed interest in a system that is both manageable for everyone and user-friendly. This includes transparent, intelligible evaluation reports that teachers can understand and incorporate into their teaching; a user-friendly software interface for streamlined uploading of documentation; and, perhaps most importantly, a streamlined observation and evaluation process to reduce the burden on administrators.

At many community input sessions groups described the great lengths that administrators had to go to in order to complete teacher evaluations. Farmington building administrators shared the concern that "administrators should not be giving up weekends regularly in order to complete evaluations." Similarly, Gallup elementary teachers noted, "Principals can be consumed by the time it takes [to complete evaluations]. Everything else goes to the wayside."

Put simply, administrators and teachers want more time observing and talking about quality instruction and less time sitting at computers entering data. Stakeholders across New Mexico articulated that in order for the state to create a system that is truly focused on teacher growth and development, the system must be administratively feasible.

Fair Appeals Process to Address Disagreements and/or Unreflective Observations

"An appeal process should be created for teachers to rebut their evaluation."
- *Community Advocates, Farmington Stakeholder Meeting*

"If you disagree with an evaluator's conclusion, you should have the right to appeal it."
- *Secondary Educators, Albuquerque Stakeholder Meeting*

"Teachers should get another chance to redo a bad evaluation or one that they feel was not their best."

- *Elementary Educators, Ruidoso Stakeholder Meeting*

Stakeholder groups in input sessions across the state articulated the importance of the new teacher evaluation system having an appeals process for teachers who feel their observation or rating does not accurately reflect their work in and out of the classroom. This process would allow teachers to, for example, redo an observation or submit further evidence and documentation to address any gaps in their initial documented practice.

At the heart of this recommendation is the belief that a teacher evaluation system should be focused on educator growth and development. To effectively support improvement, the system needs to accurately assess an educator's performance. Multiple stakeholder groups shared that a poor observation or evaluation may occur, sometimes for reasons outside an educator's control. They did not feel it fair that their overall rating be dragged down by such an outlier.

Note: Implementation of multiple observations could complement the purpose of an appeal propose. By reducing the relative significance of any single observation on a final evaluation rating, educators are more likely to view evaluations as more authentically reflecting their teaching practice. When well implemented, both components could help build trust in the system.

Access to Resources and Exemplars That Demonstrate Levels of Teaching Defined in Rubrics

"We want exemplars—videos of what an awesome teacher looks like. We should be able to see those. Having videos takes some of the tension out of evaluations."
- *Secondary Educators, Deming Stakeholder Meeting*

Several stakeholder groups expressed an interest in having access to resources that demonstrate and illustrate teaching performance levels described in observation rubrics. For example, several educator groups were interested in the idea of a bank of exemplars of instruction at the various evaluation performance levels, perhaps even drawn from video evidence of New Mexico teachers.

These resources would not only serve to create a shared understanding of effective (and ineffective) instruction among teachers and evaluators, they would also elucidate a pathway to professional growth. By viewing gradated levels of performance, teachers would have concrete

17

examples of how to improve their instruction. These videos could be further coupled with video evidence of teachers' own instruction, comparing the two to illuminate strengths and areas for improvement.

18

Appendix C:
Task Force Schedule

Date	Location
August 10 & 11, 2019	Bernalillo, NM
September 28, 2019	Bernalillo, NM
October 19, 2019	Bernalillo, NM
November 16 & 17, 2019	Pojoaque, NM
January 9, 2020	Bernalillo, NM