New Mexico Teacher Evaluation System

Technical Guide

Business Rules and Calculations

2017-2018 School Year

Teacher Summative Evaluation

Last Updated: August 2018

Contents

Overview and Purpose

The NMTEACH Educator Effectiveness system is designed to establish a framework for continuous improvement and professional growth for teachers and principals, which ultimately will promote student success. The NMTEACH system was created to ensure that every student has equitable access to an effective principal and teacher every day they are in school. Implementing a rigorous, uniform observation protocol, providing immediate constructive feedback, using meaningful student data, and other multiple measures will provide valuable information to aid the professional development and growth of each teacher and principal.

In the interest of transparency and to promote best practices for all educators, NMTEACH aims to provide information that is timely, relevant, and constructive with respect to critical areas of teacher performance. By releasing reports in the fall, it is the goal of NMTEACH to also provide information that will help to guide the discussions on planning the school year. The purpose of this document is to describe the business rules associated with NMTEACH (e.g., how summative scores are computed and combined, which teachers are included, and how measures are calculated).

As in previous evaluation cycles, there are four main performance areas that NMTEACH aims to measure: how teachers support a student's opportunity to learn, how teachers impact student achievement, the quality of the instruction in the classroom, and teacher professionalism. Each performance area corresponds to an evaluation question and one or more performance measures. The performance areas, evaluation questions, and performance measures for each are listed in the table below.

| Performance Area	Evaluation Question	Performance Measures
Student Opportunity to Learn (OTL)	To what extent do teacher practices and behaviors create student opportunity to learn?	Student/Parent OTL Survey; Domain 1, 2, and 3; Teacher Attendance
Student Achievement	To what extent does the teacher enable students to exceed expectations of achievement on standardized tests?	Student Achievement Gains
Instructional Quality	To what extent do teacher practices and behaviors maintain high standards of academic quality?	Domain 1, 2, and 3
Professionalism	To what extent do teachers contribute to positive school culture and climate?	Domain 4, Teacher Attendance

The NMTEACH Theory of Action

The flowchart on the following page maps the theory of action for the NMTEACH evaluation system, which is characterized by the relationship between Teacher Effectiveness, Instructional Practice, and Student Achievement. The Theory of Action is defined as a series of *if, then* statements that encode beliefs about a system where an arrow represents an if-then belief. At its core, the NMTEACH Theory of Action reflects the belief that if teacher effectiveness improves, then instructional practice will improve, which will then improve student achievement. In other words, Teacher Effectiveness affects what instructional decisions a teacher uses as part of his or her practice, which in turn impacts student achievement. This core belief is displayed in the center row of the Theory of Action flowchart: Teacher Effectiveness → Instructional Practice → Student Achievement. NMTEACH uses a variety of measures to assess these relationships.

Educator effectiveness is measured through the NMTEACH Observation Rubric, Multiple Measures, and the value-added score (VAS), which is the teacher contribution to student achievement. By using different types of measurements, the NMTEACH Educator Effectiveness system looks at the many different ways a teacher contributes to student learning. The details of how each of these measures work is the bulk of this technical document.

While a variety of measures can be used to assess student achievement, the assessments selected to be a part of this evaluation cycle are the following: the Standards Based Assessments (SBA), the Partnership for Assessment of Readiness for College and Careers (PARCC) assessments, end of course exams (EOCs), and Istation. These specific assessments were carefully selected to promote the idea that all students deserve the same level of high quality education. Because these assessments are easily accessible to all schools, school participation is not dependent upon such elements as geographic region, and charter school designation. All school districts in New Mexico – large and small, rural and urban, charter schools and traditional schools – have access to the materials and training necessary to administer these assessments. Also, the cost of these assessments is absorbed by the State and districts, ensuring students from all economic backgrounds are able to participate. And importantly, most schools are already administering these assessments, eliminating the need for additional testing.

In the Theory of Action, the most complex element to measure is the teacher's instructional practice—the day-to-day decision making about curriculum, lesson planning, classroom management, integrating teachable moments, and all the large and small pieces that coalesce to create effective instruction. By using a variety of measurements, the NMTEACH Educator Effectiveness system aims to capture this complexity. The overall goal of the summative report is to recognize educator quality and effectiveness, which, in turn, will guide larger discussions on replication of best practices and continued professional learning opportunities for all teachers.

The NMTEACH Theory of Action
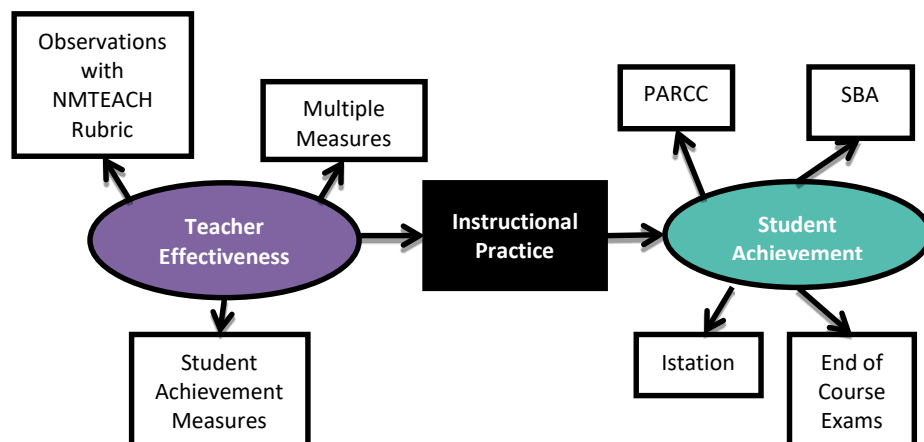
The Components of the NMTEACH Educator Effectiveness System

The NMTEACH system utilizes up to three years of data, depending upon the data submitted by the District or Charter for a given teacher. Summative reports are made up of three general components: the Observation component[1]; the Multiple Measures component; and the Student Achievement component, as shown on the theory of action flowchart. The Observation component of the evaluation refers to Domains 2 and 3 (Creating an Environment for Learning, and Teaching for Learning) of the NMTEACH Observation Rubric. Domains 1 and 4 (Preparation and Planning, and Professionalism) of that rubric are a part of the Multiple Measures component. Multiple Measures also includes teacher attendance and surveys of students and parents. The Student Achievement component of the evaluation includes different standardized student achievement measures including SBA, PARCC, EOCs, and Istation.

The NMTEACH Educator Effectiveness System

Component of NMTEACH System	Measures Used
Student Achievement Measures (STAMs)	SBA
	PARCC
	EOCs
	Istation
Observation	Domains 2 and 3
Multiple Measures	Domains 1 and 4
	Teacher Attendance
	Surveys

Based on feedback dating back to April, 2017 and formalized in August, 2017, NMTEACH transitioned to two "steps", depending upon whether or not the teacher has student achievement growth data.

---

[1] Domains 2 and 3 (Creating an Environment for Learning & Teaching for Learning) are based on in-person observations of a teacher's classroom. The evaluation of Domains 1 and 4 (Preparation and Planning & Professionalism) occurs outside of the classroom.

Step 1 includes those teachers who have no usable student achievement data in the last three years. This would include teachers who teach courses that are not tied to one of the assessments being used, and teachers who have less than ten students with reported data in the last three years.

For more information on which courses are linked to assessments being used for NMTEACH EES purposes, please refer to the Course Assessment Linkage Document for SY 2017-2018:

https://webnew.ped.state.nm.us/bureaus/educator-growth-development/nmteach/nmteach-toolbox/

Step 2 will include teachers with one to three years of student achievement data.

The Two Steps for Teacher Summative Reports

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measures - Domains 1 and 4	Multiple Measures – Teacher Attendance and Surveys
Step 1: Teachers who have no student achievement data in the last 3 years	0%	50%	40%	10%
Step 2: Teachers with 1-3 years of student achievement data	35%	40%	15%	10%

The Summative Score Calculation

The final combination of these components forms a single summative score, with a range of 0 – 200, where higher scores indicate greater teacher effect. In the most general case of no missing data (see Adjustments to Possible Points to Account for Missing Data for exceptions), the possible points are distributed as follows:

The Possible Points for Summative Score

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measures Domains 1 and 4	Multiple Measures Teacher Attendance and Surveys
Step 1: Teachers who have no student achievement data in the last 3 years	0	100	80	20
Step 2: Teachers with 3 years of student achievement data	70	80	30	20

The following sections will elaborate how points are earned toward each of these measures. Once the points earned are determined, they are added together to give the final summative score for that teacher:

$$Summative\ Score = STAMpts + D2\&3pts + D1\&4pts + Surveypts + Attendancepts$$

STAMpts are the number of points earned from available student achievement measures, *D2&3pts* are the points earned from Domains 2 and 3, *D1&4pts* are the points earned from Domains 1 and 4, *Surveypts* are the points earned from the Student or Family surveys in the Multiple Measures component, and *Attendancepts* are the points earned from the Attendance data submitted by the teacher's district for the Multiple Measures component.

A Note on Rounding: Rounding only occurs on the final summative scores, which are rounded to the nearest whole number for each category (STAM, Observation, and Multiple Measures) and then added together. Throughout this guide, some scores are rounded for ease of reading, but in the true calculations, no rounding occurs until the end.

Example 1: Calculating the Summative Score

The chart below shows an example of the table from the first page of the summative report. This teacher has data for Student Achievement, data for Surveys and Teacher Attendance, Domains 2 and 3 (Observation), and Domains 1 and 4 (Multiple Measures)

Category	Possible Points	Points Earned
Student Achievement	70.00	54.60
	70.00	54.60
Observations Domains 2&3	80.00	56.80
	80.00	56.80
Domains 1&4	30.00	25.21
Teacher Attendance	10.00	7.50
Surveys	10.00	8.35
Subtotal	50.00	41.06
Total Evaluation	200.00	152.46

The Summative Score for this teacher is 152.46, which is the sum of all points earned for all of his or her measures and then rounded to the nearest whole number:

$$Summative\ Score = STAMpts + D2\&3pts + D1\&4pts + Surveypts + Attendancepts$$

$$Summative\ Score = 54.60 + 56.80 + 25.21 + 8.35 + 7.5 = \mathbf{152.46}.$$

Initiating a Summative Report

To receive a summative evaluation, a teacher must have at least one finalized observation with sufficient data in Domains 2 & 3 to receive an observation score (e.g., 60% or more domain elements with data – see Classroom Observation score below).

Determining Points Earned for Each Summative Evaluation Component

The following section discusses in detail how the points for each of the components of the summative evaluation are calculated. These components include Classroom Observation (Domains 2 and 3), Multiple Measures (Domains 1 and 4), Student and Family Surveys, Teacher Attendance, and Student Achievement Measures (i.e., Teacher Value-Added Scores).

Classroom Observation

Similar to previous evaluation cycles, the classroom observation component of NMTEACH is composed of two domains from the NMTEACH Framework for Effective Teaching – *Domain 2: Creating an environment for learning*, and *Domain 3: Teaching for learning*. For School Year 2017-2018, only one observation was required for teachers who earned at least 146 points and at least 50% of their possible STAM points, on the previous summative evaluation. All other teachers are required to be observed a minimum of two times, depending on what plan was chosen by the district. The deadline for all observations was May 31, 2018.

Collection of Teacher Observation Data

In general, observation data are being collected through a platform currently administered by Frontline. Frontline is being used as an observation data collection tool, including a repository for those artifacts that are necessary for Domain 1 and Domain 4. In addition, Frontline enables teachers to create a Professional Development Plan and provides access to high quality professional development resources. In the current iteration, Frontline enables a site-based Administrator to manage the user base at that school site. The Frontline Administrator is required to create user profiles for each teacher to be observed

at the school site. If a teacher previously exists in Frontline at a different school site, the Administrator is to follow steps to transfer the user to their school site and ensure that the teacher is inactive at the previous site. Please note that observation data from the previous site will not always be transferred to a teacher's new location.

Through systematic and thorough data auditing, PED cleans and validates this observation data throughout the school year. This step is critical to ensure that each teacher can be matched with PED information located in the STARS, Teacher-Student Data Link (TSDL) data warehouse (Accuroster), and the other information obtained through the NMTEACH data generation processes.

Scoring Teacher Observation Data

The classroom observation component has a score range of 10 – 50 and is based on 10 separate indicators, where higher scores indicate more effective practices and behaviors. Five of the indicators are part of Domain 2, and five of the indicators are part of Domain 3. When a teacher is observed, he or she receives a score on each separate observation occasion. The score for the Observation component of the summative evaluation is calculated by, taking the mean average of the multiple observations by indicator and then computing the sum of the average indicator scores a teacher earned in Domains 2 and 3.

Example 1: Teacher A Observation Score with No Missing Data (Domains 2 & 3)

	Domain 2					Domain 3					Total	Multiplier=10/number of complete responses	Final Score = Grand Total/number of obs.	Weight = Final Score/50
	2a	2b	2c	2d	2e	3a	3b	3c	3d	3e				
Obs 1	2	3	2	4	3	4	2	3	3	4			32	0.64
Obs 2	2	3	2	4	2	4	3	3	3	5				
Mean	2	3	2	4	3	4	3	3	3	5	32	10/10 = 1		

In the previous example, the teacher earns a final score of 32, which is out of 50 points total and yields a weight of 0.64. The multiplier column allows for the potential of unobserved indicators. For the observation to be considered valid, at least 60% of the indicators must be scored. The following table models the calculation procedures for missing data. In this case, maximum scores for each observed item will be calculated as normal. Any item that is missing all data points will be considered incomplete. A multiplier will be calculated by dividing the total possible complete responses (for Domains 2 & 3, this is 10) by the total actual complete responses. This multiplier is then used to determine the final score.

Example 2: Teacher B Observation Score with Missing Data (Domains 2 & 3)

	Domain 2					Domain 3					Total	Multiplier = 10/num of complete responses	Final Score = ΣTotal * Multiplier	Weight = Final Score/50
	2a	2b	2c	2d	2e	3a	3b	3c	3d	3e				
Obs 1	2	3	■	4	■	4	■	3	3	4			32.5	.65
Obs 2	2	3	■	4	2	4	■	3	3	5		10/8 = 1.25		
Mean	2	3	■	4	2	4	■	3	3	5	26			

To compute the points earned on the summative evaluation for observations, NMTEACH multiplies the weight by the possible points for observations.

Example 3: Calculating Observation Scores: Teachers A and B

	Weight	Possible Summative Points	Summative Points Earned
Teacher A	.64	80	51.2
Teacher B	.65	100	65

Multiple Measures 1: Domain 1 and Domain 4

Similar to the classroom observation data, Domain 1 and Domain 4 data are collected in the Frontline system.

Domain 1 and Domain 4 are computed in much the same way that the Classroom Observation scores are computed. As is done with Domains 2 and 3, the maximum of the multiple observations by indicator is divided by the sum of the maximum scores a teacher earned in Domains 1 and 4. Furthermore, Domains 1 and 4 have 12 indicators or elements, thus the maximum score is 60, Domain 1 and Domain 4 must be scored at a minimum of one time per year.

Domain 1 & Domain 4	Completion and Data Available Date
2017-18 Submission Deadline	May 31, 2018

For example, suppose a teacher is scored on Domain 1 and Domain 4 during semester 1 and semester 2 and receives the following profile of scores:

Example 4: Calculating Domains 1 & 4 for Teacher A (No Missing Data)

	Domain 1						Domain 4						Total	Multiplier=12/number of complete responses	Final Score = Grand Total/number of obs.	Weight = Final Score/60
	1a	1b	1c	1d	1e	1f	4a	4b	4c	4d	4e	4f				
Sem 1	2	3	2	4	3	4	4	2	3	2	4	3			44	.7333
Sem 2	3	3	2	4	2	5	4	3	4	3	5	5				
Max	3	3	2	4	3	5	4	3	4	3	5	5	44	12/12 = 1		

In the previous example, Teacher A earned a Domain 1 and Domain 4 score of 44, which is 73% of the total possible 60 points. In the following example, the teacher is missing two indicator scores from the semester 1 scoring event and is missing one indicator score from the semester 2 scoring event. Therefore, notice how the multiplier is adjusted to account for the unscored indicators. For the Domain 1 and Domain 4 score to be considered valid, at least 60% of the indicators must be scored and finalized in the Frontline system.

Example 5: Calculating Domains 1 & 4 for Teacher B (Missing Data)

	Domain 1						Domain 4						Total	Multiplier=12/num of complete responses	Final Score = ΣTotal * Multiplier	Weight = Final Score/60
	1a	1b	1c	1d	1e	1f	4a	4b	4c	4d	4e	4f				
Sem 1	1	2	■	3	2	3	4	■	3	■	4	3				
Sem 2	2	3	2	2	3	3	4	3	3	■	5	5			39.27	.655
Max	2	3	2	3	3	3	4	3	3	■	5	5	36	12/11 = 1.0909		

To compute the points earned on the summative evaluation for Domains 1 and 4, the weight is multiplied by the possible points for Domains 1 and 4.

Example 6: Calculating Domains 1 & 4 scores for Teachers A and B

	Weight	Possible Summative Points	Summative Points Earned
Teacher A	0.733	80	58.64
Teacher B	0.655	30	19.65

Multiple Measures: Student or Family Survey

The surveys consist of 10 items that measure student or family perception of the opportunity to learn created by the teacher. Student Surveys, for students in grades 3-12, have two windows for completion. This is to account for classes that are only one semester long or courses on a block schedule. Family surveys, completed by families of students in grades K-2, have one window for completion. Student and family survey windows closed May 18, 2018.

The Student and Family Surveys are as follows:

Qnum	Question
1	My child's teacher expects my child to do his/her best.
2	My child's teacher checks that my child understands the work.
3	My child's teacher can tell me about my child's strengths and weaknesses.
4	My child's teacher includes me in helping to improve my child's r&m skills.
5	My child's teacher provides clear instructions for homework.
6	My child's teacher answers my questions.
7	My child's teacher provides regular feedback about my child's learning.
8	My child's teacher provides regular feedback about my child's behavior.
9	My child's teacher notices when something is bothering my child.
10	My child's teacher invites me to the classroom.

	DontKnow	Never	Rarely	Inconsistently	Consistently	Always
Score	Missing Info	0	1	2	3	4

Parent Survey

Qnum	Question
1	I know what I should be working on in class.
2	My teacher introduces a new topic by connecting to things I already know.
3	My teacher checks to see if I understand.
4	My teacher wants me to explain my answers.
5	My teacher knows when I understand, and when I do not.
6	My teacher explains things in different ways so I can understand.
7	My teacher wants me to try to correct my mistakes.
8	My teacher takes the time to summarize what I learn each day.
9	My teacher expects me to do my best.
10	My teacher notices when something is bothering me.

	Never	HardlyEver	Sometimes	Usually	AlmostAlways	Always
Score	0	1	2	3	4	5

Student Survey

Family Surveys are administered for teachers who teach students in grades K – 2. Student Surveys are administered for teachers who teach students in grades 3 – 12. The survey collection system does not allow incomplete surveys to be recorded. At least 10 complete surveys from unique users of one type (either family or student) must be recorded in order for the survey responses to be considered valid. If a teacher has *both* family and student surveys, then family surveys are used *only if* there are fewer than 10 student surveys and greater than 10 family surveys. Otherwise, student surveys are used. Surveys are administered through a secure system called Voice NM.

Consider the two examples: the first reflects a student survey with complete data and the second a family survey with complete data.

Example 7: Calculating Student Survey Points

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total	Final Score = Average of Total Scores	Weight = Final Score/ 50
Stu 1	3	3	3	4	3	4	4	5	5	3	37		
Stu 2	2	3	2	4	2	4	3	3	3	3	29	33.6667	0.6733
Stu 3	3	4	3	5	3	3	3	2	4	5	35		

Example 8: Calculating Family Survey Points

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total	Final Score = Average of Total Scores	Weight = Final Score/40
Par 1	2	1	3	4	3	3	3	4	4	3	30		
Par 2	2	1	2	4	2	4	3	3	3	3	27	29	0.725
Par 3	3	4	3	2	3	3	3	2	4	3	30		

To compute the points earned on the summative evaluation for surveys, NMTEACH calculates the average total score and then divides by the number of possible survey points (50 for students and 40 for parents) to get the final weight. Here are a few examples for teachers with different weights:

Example 9: Calculating Survey Score

	Weight	Possible Summative Points	Summative Points Earned
Teacher 1	0.673	10	6.73
Teacher 2	0.681	10	6.81
Teacher 3	0.725	10	7.25
Teacher 4	0.695	10	6.95

Multiple Measures: Teacher Attendance

Teacher absences are submitted to the PED by each district, through a STARS Production template. For the 2017-2018 school year, teacher attendance data will only be reported through the STARS system for the period beginning of contract year through the End of Year STARS reporting period. Absences due to Family and Medical Leave Act, bereavement, jury duty, military leave, religious leave, professional development, or coaching are excused and should not be reported as absences by the District to the PED reporting systems. Absences are reported from Beginning of Year, or Contract Date through End of Year. The reporting and verification windows are presented in the table below.

Attendance Interval	Verification Period
Beginning of Contract Year – End-of-Year	July 13,2018 (STARS)

In addition to the exceptions mentioned previously, NMTEACH allows for a grace period of up to 6 days. In other words, if the district reports 6 days absent for a teacher, the teacher still receives 100% of the possible points for the Multiple Measure: Attendance. However, if a district reports more than 6 days absent for a teacher all of the days are figured into the score. The maximum number of absences is 20, the following formula is applied to compute the weight associated with teacher attendance:

$$W_{\text{Teach Att}} = \frac{(20 - number\ of\ absences\ reported)}{20}.$$

If the teacher has 20 or more absences, then the weight is automatically equal to zero. Further, if total absences reported is less than or equal to 6, then teacher receives 100% of attendance points; if total absences reported is greater than 6, then standard formula applies. For example, consider the following reported data:

Example 10: Teacher Attendance Reported from One District

	Number of Absences Reported	Weight	Possible Summative Points	Summative Points Earned
Teacher A	0	1	10	10
Teacher C	19	0.05	10	.5
Teacher D	25	0	10	0.00
Teacher E	6.5	.675	10	6.75

If a teacher has reported absences from *multiple districts* the attendance is calculated at the last observed location.

Improved Student Achievement: STAMs (Student Achievement Measures)

All STAMs are computed using a value-added model (VAM). There are four main steps to including STAMs in the summative report:

1. Verification of course-teacher-student linkages by districts.
2. Estimation of the Value-Added Score (VAS) for each teacher by course group and year. Separate models are run for each outcome measure type. All teachers with data for each outcome, course group, and year are included in the models.
3. Calculation of Student Achievement Measure (STAM) score using a weighted average of all relevant VAS for each teacher.

This section will describe each of these steps in detail.

Step 1: Data Verification by district with the Accuroster System

Accuroster, also known as the Teacher Student Data Link (TSDL) is a web-based program that allows districts to review, verify, and update the linkage of teachers and students to courses that in turn will connect them to the appropriate Student Achievement Measure. The Accuroster is populated from a STARS report that contains roster data reported by the districts at each of the 40-, 80-, and 120-day Snapshots. The Accuroster can be found here with the proper login credentials: https://eesaccuroster.ped.state.nm.us/MainMenu.aspx

Prior to data being pulled for use in calculating summative scores, the Districts are granted access to the Accuroster and can view all of the students that are rostered within each teacher's class in the district. Only district-level users are supplied with logins and passwords. These permissions are based upon STARS reporting information.

The district-level users are able to modify what teachers are bound to the students for each course. Updates are made via batch download and upload into the system.

After all updates have been made in the Accuroster, the data table housing these updates are supplied back to evaluators and is then used to inform the mapping for STAM, which is discussed more in-depth in the following section.

Step 2: Estimation of Value-Added Scores

The Value-Added Model is used to compute a teacher's unique contribution to student achievement. A two-level mixed-effects regression model is employed to estimate the statistical parameters that isolate the unique contribution of the teacher to student achievement. Two prior data points in reading and math (when available) are collected for each student and used to predict the current year score for the student.

Depending on the STAM, this outcome score could be the student's score on SBA, PARCC, Istation (English or Spanish), or EOCs. The model adjusts for the proportion of time that a student is with a teacher using the 40, 80, and 120 day snapshots, and the model adjusts for whether or not the course is designated as an intervention class (ESL, Resource, or Reading Intervention classes). These factors are used to compute an expected current year score for the relevant outcome assessment for each student. For courses that include students in multiple grade levels, the student's grade is also included in the model as a covariate. A teacher's Value-Added Score (VAS) is based on the average of the differences between the actual student score and the predicted student score. The following defines the outcome variable and prior achievement used for each model.

Prior Achievement Required for Each Grade and Outcome Combination

Grade and Outcome	MATH Prior Achievement 1 Year ago	MATH Prior Achievement 2 Years ago	ELA Prior Achievement 1 Year ago	ELA Prior Achievement 2 Years ago
iStation Grade K (EOY K)	N/A	N/A	Istation MOY Kindergarten	Istation BOY Kindergarten
iStation Grade 1 (EOY 1st grade)	N/A	N/A	Average of Istation BOY and MOY 1st grade	Istation EOY Kindergarten
iStation Grade 2 (EOY 2nd grade)	N/A	N/A	Average of Istation BOY and MOY 2nd Grade	Istation EOY 1st grade
PARCC Grade 3 Math	N/A	N/A	3rd Grade PARCC/SBA ELA	Istation EOY 2nd Grade
PARCC Grade 3 ELA	3rd grade PARCC Math	N/A	Istation EOY 2nd Grade	N/A
SBA/PARCC Grade 4 (ELA, Math, Sci.)	3rd grade PARCC Math	N/A	3rd grade PARCC ELA	N/A
PARCC Grade 5 (ELA and Math)	4th grade PARCC Math	3rd grade PARCC Math	4th grade PARCC ELA	3rd grade PARCC ELA
PARCC Grade 6 (ELA and Math)	5th grade PARCC Math	4th grade PARCC Math	5th grade PARCC ELA	4th grade PARCC ELA
SBA/PARCC Grade 7 (ELA, Math, Science)	6th grade PARCC Math	5th grade PARCC Math	6th grade PARCC ELA	5th grade PARCC ELA
PARCC Grade 8 (ELA and Math)	7th grade PARCC Math	6th grade PARCC Math	7th grade PARCC ELA	6th grade PARCC ELA
PARCC HS Math (Algebra I, Algebra II, and Geometry)	Prior year PARCC Math	2 years prior PARCC Math	Prior year PARCC ELA	2 years prior PARCC ELA
PARCC HS ELA Grade 9	8th grade PARCC Math	7th grade PARCC Math	8th grade PARCC ELA	7th grade PARCC ELA
PARCC HS ELA Grade 10	9 th grade PARCC Math	8th grade PARCC Math	9th grade PARCC ELA	8th grade PARCC ELA

PARCC HS ELA Grade 11	10th grade PARCC Math	9th grade PARCC Math	10th grade PARCC ELA	9th grade PARCC ELA
SBA Science Grade 11	10th grade PARCC Math	9th grade PARCC Math	10th grade PARCC ELA	9th grade PARCC ELA
EOCs 4th	3rd grade PARCC Math	N/A	3rd grade PARCC ELA	N/A
EOC (grades 5-12)	Prior year PARCC Math	2 years prior PARCC Math	Prior year PARCC ELA	2 years prior PARCC ELA

Data Requirements and Structures

A teacher's VAS will be based on only those students with a complete data history that are rostered to a course linked with that teacher. This means, to be included in the VAM, students must have data associated with them in all fields. Those fields include: two years of math and reading, proportion of time with the teacher, grade level of the student, and intervention status of the course. If a student is missing one or more of these data points, they will **not** be included in the model. The only exceptions to these requirements are for early elementary grades (K-4) where students may not have two years of prior data. See below for further explanation.

In general, the student level data set to estimate the VAM is organized as follows:

Example 11: Student Level Data File for Standardized Tests (PARCC and EOC) Grades 5 – 11

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Valid Record	Student#	Teacher#	CourseGroup	Course#	zRead18	ProportionofTime18	Intervention	Z Prior Math 1	zPrior math 2	Z Prior Read 1	Z Prior Read 2	ConditionalSEM WeightRead	StudentGrade18	Student Grade Prior 1	Student Grade Prior 2
Y	123456789	123456	ELA 10	1016	1.05	100	0	0.89	1.01	0.98	1.11	0.27	10	8	7
Y	987654321	654321	Read 7	1004	1.47	75	1	1.2	0.83	1.93	1.8	0.39	7	6	5
N	199999991	123456	ELA 10	1016	-0.22	25	0					0.31	10		
N	911111119	654321	Read 7	1004	0.58	75	1	1.2		1.93		0.29	7	6	

The data table above illustrates four hypothetical students assigned to two different teachers. Each row represents one student. Two of the students are assigned to a teacher of ELA 10, and two students are assigned to a teacher of reading Grade7. The first column indicates whether the case is valid to use in the model; the second column lists the Student ID; the third column lists the Teacher Number; the fourth column lists the Course Group; the fifth column lists the 4-digit State Course Number; the sixth column lists the standardized student test scores in the current year; the seventh column lists the proportion of time that the student is with the teacher (40-, 80-, and 120-day), the eighth column indicates the

classroom setting (intervention or not), the ninth column lists the student's standardized test score in math from the prior year, the tenth column lists the student's standardized test score in math from two years prior, the eleventh column lists the student's standardized test score in reading from the prior year, the twelfth column lists the student's standardized test score in reading from two years prior, the thirteenth column lists the conditional standard error weight for reading, the fourteenth column lists the student's grade level in the current year, the fifteenth column lists the student's grade level from the prior year, and the sixteenth column lists the student's grade level from two years ago.

As displayed in the previous table, if a student is missing any prior year data, then they will not be included in the Value-Added Model. One of the students assigned to the ELA 10 teacher is missing all prior year data, while one of the reading Grade 7 students is missing PARCC data from 2017; in both cases they represent invalid cases, as they are missing data. The actual dataset has a row for each student in the course group.

Example 12: Teacher-Level Data File

Teacher Number	N (Number)	Grand Mean Centered zRead17	Grand Mean Centered zMath17
123456	27	0.45	0.38

The table above illustrates the teacher level file for ELA 10; though it can be generalized to other grade levels, tests, and course groups. The first column lists the Teacher Number; the second column lists the total number of students associated with the teacher some of which are valid because they have prior achievement data and some which are invalid because they do not have prior achievement data; the third column lists the average standardized reading scale score from the prior year, which has been grand mean centered by course group; and the fourth column lists the average standardized math scale score from the prior year, which has also been grand mean centered by course group. The actual data set has a row for each teacher associated with the course group. The general statistical model is presented below:

$$NSY_{ij}^T = \pi_{0j} + \pi_1(NSSM)_{ij}^{T-1} + \pi_2(NSSM)_{ij}^{T-2} + \pi_3(NSSR)_{ij}^{T-1} + \pi_4(NSSR)_{ij}^{T-2}$$
$$+ \pi_5(Intervention)_{ij}^T + \pi_6(Proportion)_{ij}^T + e_{ij}^T$$
$$\pi_{0j} = \theta_{00} + \theta_{01}(\overline{NSSM_j} - \overline{NSSM_{..}})^{T-1} + \theta_{02}(\overline{NSSR_j} - \overline{NSSR_{..}})^{T-1} + r_{0j}$$

Student-Level Variables

Student-Level Variables	Description
NSY_{ij}^{T}	A standardized PARCC (Math, Reading), SBA (Science), or EoC scale score for the ith student associated with the jth teacher in year = T (current year). For the 2016 VAS, T equals 2018.
$(NSSM)_{ij}^{T-1}$	A standardized PARCC/SBA Math achievement score from the prior year for the ith student associated with the jth teacher
$(NSSM)_{ij}^{T-2}$	A standardized PARCC/SBA Math achievement score from two years prior for the ith student associated with the jth teacher
$(NSSR)_{ij}^{T-1}$	A standardized PARCC/SBA Reading achievement score from the prior year for the ith student associated with the jth teacher
$(NSSR)_{ij}^{T-2}$	A standardized PARCC/SBA Reading achievement score from two years prior for the ith student associated with the jth teacher
$(Intervention)_{ij}^{T}$	A [0,1] variable that indicates whether or not teacher j's course during time T in which that student i was enrolled was an intervention course
$(Proportion)_{ij}^{T}$	A variable that reflects the amount of time in year T the ith student was enrolled in the jth teacher's course (as reported in the 40, 80, and 120 day snapshots)

Teacher-Level Variables

Teacher-level variables	Description
$(\overline{NSSM}_j - \overline{NSSM}_{..})^{T-1}$	Average standardized prior year PARCC Math achievement score of students associated with the jth teacher that has been grand mean centered relative to the course group
$(\overline{NSSR}_j - \overline{NSSR}_{..})^{T-1}$	Average standardized prior year PARCC/SBA Reading achievement score of students associated with the jth teacher that has been grand mean centered relative to the course group

Student-Level Coefficients

Student-Level Coefficients	Description
π_{0j}	Is the adjusted mean achievement for teacher j after controlling for prior academic preparation, intervention status, and proportion of time the student was with the teacher
π_1	Is the fixed student level effect of PARCC math achievement from the prior year
π_2	Is the fixed student level effect of PARCC math achievement from two years prior
π_3	Is the fixed student level effect of PARCC/SBA reading achievement from the prior year
π_4	Is the fixed student level effect of PARCC/SBA reading achievement from two years prior
π_5	Is the fixed student level effect of whether the class was intervention or not
π_6	Is the fixed student level effect of the proportion of time the student was assigned to a teacher
e_{ij}^T	Is the unique contribution associated with the ith student in the jth teachers at time = T

Teacher-Level Coefficients

Teacher-Level Coefficients	Description
θ_{00}	Is the adjusted mean Test achievement in the course group, across the State
θ_{01}	Is the fixed teacher level effect of the teacher level average PARCC math achievement from the prior year (grand mean centered)
θ_{02}	Is the fixed teacher level effect of the teacher level average PARCC/SBA reading achievement from the prior year (grand mean centered)
r_{0j}	Is the unique contribution of teacher j at time = T

Rules of Thumb for Interpreting VAS

- If VAS is equal to 0, then the teacher's students made, on average, the growth expected, when compared to their peers of similar academic history
- If VAS is greater than 0, then the teacher's students made, on average, more than the growth expected, when compared to their peers of similar academic history.
- If VAS is less than 0, then the teacher's students made, on average, less than the growth expected, when compared to their peers of similar academic history.

Student Achievement Measure Options

STAM	Description
Istation	Early literacy test K, 1 and 2
EOC	End of Course exams
PARCC	Partnership for Assessment of Readiness for College and Careers
SBA	Standards Based Assessment in Science (grades 4, 7, 11)

Step 3: Teacher VAS to Be Included in Summative Report Score

All step 2 teachers that receive a summative report will have a VAS score calculated for every year and course group that they taught. All valid VAS scores calculated will be used in the STAM Score for the summative report.

If a teacher was a substitute in one or more years in which they have VAS, these data points will not be included.

Step 4a: Combining VAS into Overall VAS

The overall VAS score is based on a weighted average of all of a teacher's individual course group VAS scores (as determined in the previous subsection). This means that the course groups that have more students attached to them are given a larger weight than course groups with fewer students. Adding weight to these scores occurs for two reasons. The first is a statistical reason: the reliability of the value-added scores is stronger when more students are used to create the estimate of how well the teacher has taught the material; it is more probable that the average score of many students is more accurate and predictive than is the average score of a small number of students. The second reason is more philosophical: when teachers impact a greater number of students, the VAS for a course group with more students should receive more weight.

The weighted average formula for the overall VAS is as follows:

$$Overall\ VAS = \frac{sum(N_k * VAS_k)}{sum(N_k)} \ for\ all\ k\ VAS\ course\ groups\ for\ the\ teacher$$

Example 13: Finding the Overall VAS

Teacher A is a high school teacher English teacher. If he had three course groups, 8th Grade ELA, 10th grade ELA, and 11th grade ELA, his data, broken out by course group and year, might look like this:

	2016		2017		2018	
Course Group	Number of Students	VAS	Number of Students	VAS	Number of Students	VAS
ELA8					10	-0.189
ELA10			25	0.296	15	0.637
ELA11			16	0.056	20	1.234

Teacher A had no PARCC data in 2016. He taught two course groups and therefore received VAS for two course groups in 2017 (ELA10 and ELA11) and for 3 course groups in 2018 (ELA8, ELA10, ELA11). To combine all of these VAS scores into the overall VAS, NMTEACH uses the following formula:

$$Overall\ VAS = \frac{25*0.296 + 16*0.056 + 10*(-.189) + 15*0.637 + 20*1.234}{25 + 16 + 10 + 15 + 20} = \frac{40.641}{86}$$
$$= 0.4726$$

Step 4b: Converting Overall VAS into Percentile and STAM Points

Because of the way individual VAS are distributed, the overall VAS is converted into a percentile using the cumulative density function (CDF) of a standard normal distribution. The percentile is then used to weight the possible STAM points to find the STAM points earned.

$$STAM\ weight = VAS\ Percentile = \Phi(VAS)$$

Where Φ is the cumulative density function for the normal distribution.

This conversion function looks like this:

Cumulative Density Function
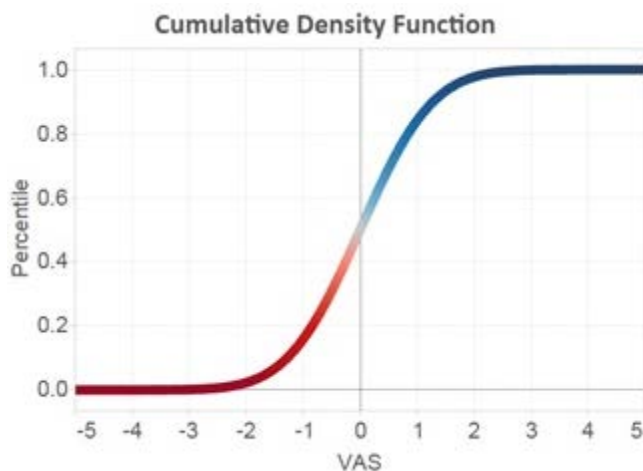
The value-added score is on the x-axis (horizontal) and the percentile is shown on the y-axis (vertical). The percentile for a particular VAS is the height of the curve at the location of the VAS on the x-axis. The percentile will always fall between 0 and 1. In Excel, the function "=NORMSDIST()" will provide the percentile for the value-added score.

To compute the points earned on the summative evaluation for STAM, the weight/percentile is multiplied by the possible points for STAM.

Example 14: Converting STAM Points Earned

	Overall VAS	Percentile =NORMSDIST(Overall VAS)	Possible Points	Summative Points Earned
Teacher 1	-0.7546	0.230	50	11.26
Teacher 2	0.00	0.5	100	50.00
Teacher 3	2.360	0.9909	100	99.09
Teacher 4	-1.895	0.03	70	2.03
Teacher 5	.500	0.69	30	20.74

Adjustments to Possible Points to Account for Missing Data

For several reasons, teachers might be missing data for one or more components of their report. Teachers who are new to New Mexico may not have student achievement data yet. Other teachers may be missing student achievement data for some years because they were either not teaching a subject with a relevant test or they were not teaching at all in those years. Also, the teachers may not have the minimum number of students with the required full history of data points. Teachers may also be missing multiple measures data. They could be missing Domains 1 and 4 because their school or district did not complete or finalize the observation in the Frontline system, did not submit the data, or did not submit the data on time. They could also have missing data if less than 60% of the Domain 1 and 4 elements were completed. Teachers could be missing attendance data if districts did not report that data. They could be missing survey data if fewer than 10 students or familes responded to the survey. The only component that will never be missing data is Observations, Domains 2 and 3. A teacher without these data will not receive an evaluation.

When data from one or more components are missing, the possible points that would have been allocated to that component are redistributed among the other components so the teacher is not penalized for missing data. This redistribution occurs using the rules laid out in the following section, Pro-Rating Points.

Pro-Rating Points

If data are missing in the multiple measures section, the possible points will be shifted to other measures. Note: *If a teacher does not have valid observations for Domains 2 and 3, he or she will not receive a summative report.* If a teacher is missing achievement data, he or she will be assigned to one of the two steps, discussed in earlier sections.

The redistribution of points is as follows:

1. A teacher will not receive a summative report if he or she does not have Domains 2 and 3.
2. If a teacher is missing Domains 1and 4, points will be transferred to Domains 2 and 3.
3. If a teacher is missing both Attendance and/or Surveys, points will be transferred to Domains 1 and 4.
4. If a teacher is missing Attendance, Surveys and Domains 1 and 4, all points for Multiple Measures will be transferred to Domains 2 and 3.

For example, a teacher in Step 2 would normally receive the following possible point distributions if all of his or her data are present. If a teacher has Surveys and Attendance, each will account for 10 points of the total 20. If one is missing, the remaining measure itself will be worth 10 points and the other 10 points will be redistributed to domains 1 and 4.

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measure - Domains 1 and 4	Multiple Measure – Teacher Attendance and Surveys
Step 2: Teachers with 1-3 years of student achievement data	70	80	30	20

If this teacher is missing Domains 1 and 4, the possible point distribution would be as follows:

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measure - Domains 1 and 4	Multiple Measure – Teacher Attendance and Surveys
Step 2: Teachers with 1-3 years of student achievement data	70	110	0	20

Likewise, a teacher missing Attendance and Surveys will have the possible point distribution seen below:

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measure - Domains 1 and 4	Multiple Measure – Teacher Attendance and Surveys
Step 2: Teachers with 1-3 years of student achievement data	70	80	50	0

In addition, a teacher missing surveys but having attendance will have the following possible point distribution:

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measure - Domains 1 and 4	Multiple Measure – Teacher Attendance and Surveys
Step 2: Teachers with 1-3 years of student achievement data	70	80	40	10

Finally, a teacher missing all multiple measures would have the following possible point distribution:

	Student Achievement	Classroom Observation - Domains 2 and 3	Multiple Measure - Domains 1 and 4	Multiple Measure – Teacher Attendance and Surveys
Step 2: Teachers with 1-3 years of student achievement data	70	130	0	0

Summary

Using the steps provided in this technical guide, an overall summative score is calculated for each teacher. The steps are as follows:

1) Determine Step.
2) Verify teacher-student-course link through the Accuroster.
3) Collect and attach assessment data to students, teachers and courses.
4) Calculate value-added scores.
5) Calculate observation and multiple measure scores.
6) Combine all scores together into summative score.

Appendix A: Glossary

Accuroster   The system for verifying and updating teacher, student, and course linkages

EES   Educator Effectiveness System

Mean   The average score

Mode   The most frequently occurring score

STAM   Student Achievement Measure - the assessment data submitted by the district for the improved student achievement portion of the summative evaluation

VAM   Value-Added Model - the multilevel mixed-effects regression model used to calculate value added scores

VAS   Value-Added Score - the score received by the teacher estimating the growth that his or her students made

Appendix B: Resources/Links

NMTEACH Website https://webnew.ped.state.nm.us/bureaus/educator-growth-development/nmteach/

NMTEACH FAQs https://webnew.ped.state.nm.us/bureaus/educator-growth-development/nmteach/frequently-asked-questions/

Domain 1 Rubric https://webnew.ped.state.nm.us/wp-content/uploads/2018/01/NMTEACH-Observation-Rubric-Domain-1.pdf

Domain 2 Rubric https://webnew.ped.state.nm.us/wp-content/uploads/2018/01/NMTEACH-Observation-Rubric-Domain-2.pdf

Domain 3 Rubric https://webnew.ped.state.nm.us/wp-content/uploads/2018/01/NMTEACH-Observation-Rubric-Domain-3.pdf

Domain 4 Rubric https://webnew.ped.state.nm.us/wp-content/uploads/2018/01/NMTEACH-Observation-Rubric-Domain-4.pdf

NMTEACH Calendar

Accuroster https://eesaccuroster.ped.state.nm.us/MainMenu.aspx

STARS Manual https://webnew.ped.state.nm.us/bureaus/information-technology/stars/

Note: All links are current as of August 2018.

Appendix C: Contacts

NMTEACH Contacts:

Policy	Seana Flanagan	seana.flanagan@state.nm.us
Data Analytics	Suchint Sarangarm	Suchint.Sarangarm@state.nm.us
Frontline	Jennifer Fresquez	jennifer.fresquez@state.nm.us
NMTEACH Liaisons	Regina Madrid	reginaA.madrid2@state.nm.us
	Jennifer Fresquez	jennifer.fresquez@state.nm.us
	Anna Delay	annaj.delay@state.nm.us
	Cherise Urioste	cherise.urioste@state.nm.us
	Mark Curran	mark.curran@state.nm.us
	Candace Green	candace.green@state.nm.us
Licensure	Rebecca Reyes	rebecca.reyes@state.nm.us