



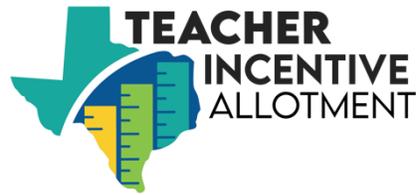
Directions for Step 2 Data Validation

Cohorts A, B, & C

Year 3 – 2021 Data Submission

Table of Contents

	Section	Page Number
A.	Preparing and Submitting Files for Step 2 Data Validation	
	Preparation of Data Files	1-2
	Submission of Data Files	3
	Common Errors Made in Step 2 Data Entry	3-4
B.	Description of Step 2 Data Validation	
	Reader-Friendly Version of Validation Checks	6
	Validation Rubric for TIA Step 2 Data Validation	10
	Statistical Analysis Protocols	13



Greetings LEA Colleagues,

We are very happy you have elected to participate in the Teacher Incentive Allotment program. This document contains detailed directions for submission of files for Step 2 Data Validation along with resources intended to provide a “big picture” perspective about what checks are performed by Texas Tech on the data submitted by districts.

We recommend that you review this document first before diving into the data templates, and then keep it handy as the data files are being prepared. If after reviewing, questions arise about the preparation of files, please contact TIA@ttu.edu.

Sincerely,

Texas Tech Team for TIA

A. Preparing and Submitting Files for Step 2 Data Validation

Preparation of Data Files

The Step 2 Data Validation process involves preparation of three (3) files:

- File 1 Designated Teacher Information (2021-2022).xlsx
- File 2 All Eligible Teacher Observation and Student Growth (2020-2021).xlsx
- File 3 Supplemental-Information-Fall-2021.docx

Download and complete the following files using Microsoft Excel version 2007 or later (for Files #1, and #2) and Microsoft Word 2007 or later (for File #3). The templates include detailed instructions for each data field. Please read the instructions carefully to avoid data entry errors and do not change any field formatting/ field names.

Deadline November 1, 2021.

Contact tia@ttu.edu

File 1 Designated Teacher Information (2021-2022).xlsx

To be completed only for teachers being submitted for a new or higher designation (Recognized, Exemplary, and Master) in SY2021-2022. Teachers must be currently employed by the district as a teacher in the 2021-2022 school year.

This Excel template is where districts will enter new teacher designations. The template contains two tabs (instructions, data entry). The first tab in the file contains detailed directions about the required information to be entered, such as district name, teacher, campus, certification role etc.

The second tab in File 1 is the data sheet where districts enter information about the teachers who will be designated in SY2021-2022. Please note that if an error is detected when entering data, the specific cell (or field) in the data sheet will be highlighted in yellow.

File 2 All Eligible Teacher Observation and Student Growth (2020-2021).xlsx

If the district used a teacher observation rubric other than T-TESS, Danielson, Marzano or NIET-TAP please contact TTU at tia@ttu.edu for a customized File 2 Teacher Observation and Student Growth Data file.

- Complete the Weighting tab
- Input teacher observation and student growth data from the 2020-2021 school year
- To be completed for all teachers in eligible teaching assignments/campuses
- Include teachers who may no longer be employed by the district
- Add only one entry per teacher

This template is where data from the “data capture year” (SY2020-2021) will be entered. This data includes several types of identifiable information along with observation and growth scores from all the district’s teachers in eligible groups (including those to be designated). This template has three tabs (weighting, instructions, data entry). Like the File 1 template, when errors are detected during data entry, the field will become highlighted in yellow indicating an error that should be corrected.

A majority of districts in Texas use T-TESS, but there are a significant number that use a pre-approved rubric (i.e., Danielson, NIET-TAP). Please note there are different File 2 templates for T-TESS, Danielson and NIET-TAP users. For LEAs using district-specific rubrics, a customize File 2 can be obtained by emailing a request to TIA@ttu.edu beginning June 14, 2021. In the email message, indicate which rubric your district uses, and we’ll send one over right away.

All the custom File 2s will have the same detailed instructions for completing the data submission portion of the file; and all have the same requirement about observation scores for designated teachers (only), including that a minimum score reflecting “proficiency” be recorded for all indicators included in the rubric crosswalk.

All File 2 versions include a tab to report weighting

Finally, all File 2s (no matter which version) contain a tab titled Weighting. This tab requests information about the different eligible teacher groups and the weighting assigned to observation, student growth and any other component included in designation decisions. This information may be found in the district’s “weighting” tab from the approved system application.

File 3 Supplemental-Information-Fall-2021.docx

In this file, please answer all questions. If questions do not apply to your LEA, please enter n/a.

Submission of Data Files

These directions provide details about how to submit your files to Texas Tech University. Submission is accomplished using a Secure File Transfer Protocol (SFTP) in order to ensure that sensitive data arrives securely for analysis. The directions also contain an outline of what data is entered into the Excel templates. Above all else, please do not email data files directly to Texas Tech, but instead use the SFTP that will be requested once data files are ready to submit.

When ready, the district submission should include:

- File 1 Designated Teacher Information (2021-2022).xlsx
- File 2 All Eligible Teacher Observation and Student Growth (2020-2021).xlsx
- File 3 Supplemental-Information-Fall-2021.docx

When data entry is complete

Compress (zip) all the files into one folder and name the compressed file “DistrictNameISD_Year” (e.g. FlowerISD_2021.zip).

Upload the compressed file using the Secure File Transfer Protocol (SFTP) link. Do NOT email data files. Emailing data files would result in a data security breach. In the event of incomplete submission, TTU will contact the district to arrange for the transfer of missing information. The district will then have **one week** to resubmit the missing data.

For technical assistance with data submission, please contact Fazil Mohammed at tia@ttu.edu. TEA and TTU will hold office hours on Monday, October 4, and October 18, from 2:00-3:00 PM. Please click [HERE](#) to register.

Deadline SFTP access may be requested as early as October 4, 2021, but will expire after seven (7) days. Prior to requesting SFTP access, please be sure all can be assembled within the 7-day window. All files for Step 2 Data Validation must be received via SFTP no later than November 1, 2021.

Contact For access to the SFTP, please contact tia@ttu.edu

Common Errors Made in Step 2 Data Entry

These common errors are ones made frequently by those submitting data during the initial years of the TIA. All of the errors can be corrected, but it may require time-consuming “back and forth” between district and TTU to resolve. So, we’re hoping that making everyone aware of the potential errors we see often might help them to be avoided. This advice includes mistakes to watch for when entering data into File 1 and useful tips for entering data in File 2.

One of the most important errors to avoid pertains to observation scores. For districts using T-TESS, each of the dimension scores for any designated teachers should be equal to or greater than “3,” which reflects a minimum requirement of “proficient,” in all of the reported dimensions. This requirement is very important because designated teachers with dimension scores lower than 3 will be automatically precluded for designation even if the system (as a whole) is approved. For districts using an alternative rubric, the minimum average score across dimensions must be equal to or greater than a “proficient” or an “at expectation” level.

TIA Data Submission: Checks to Avoid Data-Entry Errors

File 1: *For each teacher:*

- Enter the same value for “Unique ID of the educator” and “TEA ID of the educator” fields in both file #1 and file #2.
- Enter a valid County-District number (CDN) under “LEA’s County-District Number (CDN)” field.
- Enter a valid County-District Campus number (CDCN) under “Campus CDCN1” field.
- The CDCN should represent the teachers’ current school-year’s campus. No duplicate entries. Teachers designated the previous year(s) should not be listed unless they are being submitted for a higher designation. In this case, only use one entry for the higher designation. See additional directions pertaining to teachers in this circumstance in File 4 guidance (below).

File 2: *For all teachers, ensure the following entries are valid:*

- Grade Level under “Main grade level(s) for observation and student growth” field.
- Service IDs under “Main Service ID(s) for observation and student growth” field. Make sure the service ID reflects a course in an eligible teaching assignment. Do not include service IDs for courses not related to the teacher’s observation and student growth score.
- Content Area under “Main Content Area(s) for Observation and Growth” field. If the teacher is being designated in file #1, make sure the course is an eligible teaching assignment.
- Eligible teacher group Category number under the field “Eligible Teacher group.”
- Verify that the campus name under “Campus Name” field matches the campus number under the “Campus CDCN” field.
- Ensure that the CDCN is the teacher’s campus during the data-capture year (SY2020-2021).
- Make sure there are no duplicate entries. There should be only one line of data per educator.
- Any teacher who has been designated (Recognized, Exemplary, Master) must have a score of 3 (at expectation) or greater for each of the T-TESS dimensions in domains 2 and 3 in order to be approved for designation.

For all designated teachers (those in both file 1 and file 2):

- Ensure the same value was entered for “Unique ID of the educator” and “TEA ID of the educator” fields in both file #1 and file #2.
- Check that all designated teachers have met the observation performance standard of “proficient” or higher. For T-TESS, that means there should be a score of 3 or higher reported for each dimension. If the district uses an alternative rubric (other than T-TESS), ensure that the average of scores reported for each dimension reflects a level of “proficient” or higher. On the alternative rubrics, “proficiency” and “at expectation” are usually set at 3 on a 4-point scale, though there are differences among rubrics. It is incumbent upon those entering data for the district to verify that this requirement has been met.
- Ensure the Service ID entered under the field “Main Service ID(s) for observation and student growth” is present in the district eligible Service ID list.
- Ensure that the grade level and content area/course is an eligible teaching assignment.
- Ensure the teacher is still employed by the LEA as a teacher. The value for the field “Still employed by LEA” should be Y.

B. Here's What Happens to District Data in Step 2 Data Validation

This description of the Validation checks contains three sections. The first is a reader-friendly description of the statistics used in each check and a short rationale about what type of evidence each check provides toward verifying a district's system for designation. This version is intended to boost understanding of how the checks provide evidence of reliability and validity culminating in a judgment about the effectiveness of a district's system to identify the state's most effective teachers.

The second section contains the Validation scoring rubric. This rubric is used to summarize evidence, assign points for each check and ultimately make a decision about whether sufficient evidence exists to support the conclusion that a district's designation system will result in teacher selections that are valid. A district's system needs not score perfectly on every dimension, but scores on the individual components, taken together, help TEA make a decision about a district's system.

For those interested in a deeper dive into the Validation checks, the third section contains additional explanations about each of the statistical procedures, test statistics and decision rules for assigning points on the rubric.

It is important to note that due to the lack of statewide assessment data from the 2019-20 school year 2020-21 VAM results will consider students' prior achievement through the 2018-19 school year. The expectation for students will be based on the observed 2020-21 results. As a result, even if statewide achievement declines between 2018-19 and 2020-21 this will not mean that far fewer students will perform at or above expectations based on the VAM model.

Reader-Friendly Version of Validation Checks

The analyses described below are intended to verify the district designation system by comparing designations with external data and performing internal consistency checks. The purpose is to confirm that the district system functions in a manner that meets certain reliability (consistency) and validity (accuracy) standards, not to check the accuracy of an individual teacher's designation. Meeting these standards allows stakeholders to have confidence that the designation system is fair and accurate.

<p>A. Correlation between teacher observation ratings and student performance ratings</p>
<p>Check 1 The correlation coefficient between observation and growth among all eligible teachers is within the range of expected magnitude reported in the research literature.</p>
<p>For this check, analysts calculate the correlation coefficient (Person product-moment correlation) between teacher observation scores and student growth scores submitted by the district. This analysis involves looking for a trend or pattern in the relation between teaching proficiency (i.e., observation scores) and the learning exhibited by students (i.e., student growth). Based upon findings reported in peer-reviewed research literature, the expectation is that the trend or relation will be at least minimally positive. For example, the analysts will expect to see that teachers who are assigned higher observation ratings by appraisers will also have students that exhibit greater growth. Conversely, teachers who are assigned lower observation ratings by an appraiser would be expected to have students who exhibited less growth. Results from this analysis provides one piece of evidence about the validity of the designation system.</p>
<p>B. Confirm relation between district designations and student growth calculations</p>
<p>Check 2 District's designation of Recognized, Exemplary and Master (REM) teachers is found in similar proportion to historical patterns for the district in the distribution of statewide STAAR achievement.</p>
<p>For this check, analysts calculate a rank correlation coefficient (Kendall's Tau) between the designation category assigned to a teacher by the district and the equivalent category derived from state-level value-added scores calculated for teachers in the district. This analysis uses only district data from SY2020-2021 that is restricted to the group of teachers whom the district has designated and for whom a state-level STAAR-based value-added score can be calculated. This analysis looks at the <u>rank</u> of the designation (REM) and compares it to the <u>rank</u> derived from a value-added score. The expectation is that teachers assigned a Master designation would have a higher-ranking value-added score than those with an Exemplary and that teachers with this designation would have a higher-ranking value-added score than those with a Recognized designation. Results from this analysis provide another piece of evidence about the validity of the designation system.</p>
<p>Check 3 Districts accurately identify teachers for designation based on statewide VAM.</p>
<p>For this check, where possible, analysts match a district's designated teachers with their corresponding STAAR-based value-added designation from the 2020-2021 school year. Once matched, analysts compare the percentage of agreement between the district's designation and value-added designation. The expectation is that there will be a high degree of match between the district decision and value-added results. Comparisons will be made across all designation categories (REM) to determine the percentage of correct designations. Districts making correct designations for at least</p>

80% of teachers will receive full points for this check, while lower percentages will earn lower points. Districts with fewer than 60% of correct designations will receive no points for on this Validation check. Results from this analysis provides another piece of evidence about the validity of the designation system.

Check 4

Districts have a low proportion of over-identification of teachers for designation (REM).

This check uses the matched data set from Check 3, but instead examines the frequency with which districts designate teachers whose value-added scores contra-indicate designation. In other words, this check involves counting the number of times a district makes a designation, but value-added scores suggest that their designation should have been lower. The expectation is that the frequency with which a district over identifies will be low. Districts that over identify fewer than 10% of teachers will receive full points, while higher percentages of over identification will receive fewer points. Districts with greater than 25% of over identifications will receive no points on this Validation check. Results from this analysis provide another piece of evidence about the validity of the designation system.

C. Degree of inter-rater reliability for observation and growth judgements

Check 5

Across campuses, observation scores are similar for teachers in REM groups.

For this check, analysts use an analysis of variance (ANOVA) to calculate the extent to which there are similarities in observation scores for REM teachers across campuses. The expectation is that there will be small, statistically non-significant differences between the same designation levels across campuses within the district. That is, observation scores for teachers designated at the Master level are expected to be comparable regardless of campus where they teach. Similar analyses are performed for the observation scores associated with teachers in the other designation groups. If the expected level of consistency is found in the observation data, it provides evidence about the reliability of the district's designation system.

Check 6

Across campuses, percentages of student growth are similar for teachers in REM groups.

This check is similar to Check 5 in that analysis of variance (ANOVA) is used to calculate similarities for REM teachers across campuses. In this case, however, analysts are interested in comparing student growth scores, or the percentage of students who meet or exceed learning expectations. As above, the expectation is that there will be small, statistically non-significant differences between the same designation levels across campuses within the district. That is, growth scores associated with teachers designated at the Master level are expected to be comparable regardless of campus where they teach; and similar analyses are performed for the growth scores associated with teachers in the other designation groups. If the expected level of consistency is found in student-growth data, it provides evidence about the reliability of the district's designation system.

Check 7

Across assignments, observation scores are similar for teachers in REM groups.

This check is also similar to Check 5 but instead of making comparisons across campuses, this check looks for similarities in observation ratings within designation groups across teaching assignment. As before, analysis of variance (ANOVA) is used to calculate similarities among designation groups based

on teaching assignment. “Teaching assignment” and the subsequent comparisons will be defined in one of two ways based upon what data is provided by the district.

- First, assignment may mean looking at similarities in observation scores across eligible teacher groups as identified in the district TIA application; or if districts identify only one group of eligibility, then
- Second, assignment may mean looking at similarities in observation scores across teachers in STAAR-tested vs. non STAAR-tested assignments (e.g., Grade 3 math vs. Grade 5 science).

If the expected level of consistency is found in observation data across assignments, it provides evidence about the reliability of the district’s designation system.

Check 8

Across assignments, percentages of student growth are similar for teachers in REM groups.

This check is similar to Check 7, but instead of making comparisons in observation, the comparison is of student growth (percentage of students who meet or exceed learning expectations) within the designation groups across teaching assignment. As before, teaching assignment will be defined as eligible teacher groups or STAAR-tested vs. non-STAAR-tested, depending on the district system and the eligible teacher groups put forward for designation. This check is the last of four checks that are intended to provide evidence about the reliability of the district’s designation system.

D. Comparison of district designation percentage to overall statewide performance

Check 9

Percentage of students who meet or exceed expected growth in the district is approximately equal to the statewide performance standard for student growth among REM teachers.

Check 10

Observation ratings for REM teachers in the district are approximately equal to the statewide performance standards for teaching proficiency.

Both of these checks involve simple comparisons between statewide performance standards for each designation level and district-level results. Performance standards are calculated for both student growth and teacher observation ratings for the top 33% (Recognized level), top 20% (Exemplary level) and top 5% (Master level).

Performance standards for student growth are set for each designation level (i.e., Recognized = 55%; Exemplary = 60%; Master = 70%). The district’s results for the percentage of students who meet or exceed growth are compared to the performance standards. District results that meet the designation performance standards from the state are considered to be a match, but those that fall below the state standard are considered to be a mis-match.

Performance standards for teacher observation are also set for each designation level based on the average number of points assigned by appraisers for Domain 2 and 3 of T-TESS (i.e., Recognized = 3.7 points; Exemplary = 3.9 points; and Master = 4.5 points). The district’s results for appraiser ratings are compared to the performance standards. District point values that meet the performance standards are considered to be a match, but those that fall below are considered to be a mis-match. In cases where districts use an observation other than T-TESS, a crosswalk between the rubrics is performed

and equivalent levels are set (i.e., Recognized = 74% of possible points; Exemplary = 78% of possible points; Master = 90% of possible points).

The scoring criteria for these analyses will be based on the number of designation groups with which district data matches the performance standard for growth and observation scores. Greater number of points will be awarded when there is a greater number of groups with which district designations match the levels described above. Results from these analyses provide evidence about the validity of the designation system.

Check 11

The proportion of teachers at each designation level in the district's campuses are similar to other district's campuses with the same Domain 2A category.

The purpose of this check is to examine patterns in designation groups and compare them to district campuses with the same Domain 2A ratings. The expectation is that a district's proportion of designated teachers across all campus with a specific rating will be similar to other campuses across the state with the same rating. Results from analyses provide evidence about how well districts have calibrated their system to state standards as well as outcomes found among similarly-rated peer districts across the state. Results from this analysis also provides evidence about the validity of the designation system.

The scoring criteria reflect the size of the difference between the proportion of teachers designated by the district, and the proportion of designated teachers found in statewide averages of districts with the same Domain 2A ratings. Smaller differences in proportion (i.e., less than or equal to .10 difference) earn more points. Districts with proportions that differ from the statewide average by more than .50 receive "0" points on this check.

Validation Rubric for TIA Step 2 Data Validation

A. Correlation between teacher observation ratings and student performance ratings

This check is intended to confirm that teachers' appraisal scores are related to student growth scores.

Most evidence supports the accuracy of judgements Score of 3	Some evidence points toward the accuracy of judgements Score of 2	Limited evidence supports the accuracy of judgements Score of 1		None or almost no evidence supports judgements Score of 0	
1. The correlation coefficient between observation and growth among all <i>eligible</i> teachers is within the range of expected magnitude reported in research literature.		$R \geq .24$ Score of 3	$r \geq .16$ Score of 2	$r \geq .08$ Score of 1	$r < .08$ Score of 0

B. Confirm relation between district designations and VAM

These checks are intended to confirm that district designations are aligned with state-level student-growth calculations. For the current year, this analysis compares district designations to SY2020-2021 VAM data.

Most evidence supports the accuracy of judgements Score of 3	Some evidence points toward the accuracy of judgements Score of 2	Limited evidence supports the accuracy of judgements Score of 1		None or almost no evidence supports judgements Score of 0	
2. District frequency of teacher designations (i.e., <i>Recognized, Exemplary, Master</i>) are found in similar proportion to historical statewide VAM.		$T \geq 0.50$ Score of 3	$\tau \geq 0.30$ Score of 2	$\tau \geq 0.10$ Score of 1	$\tau < 0.10$ Score of 0
3. Districts <i>accurately</i> identify teachers for designation based on statewide VAM (REM).		$\geq 80\%$ Score of 3	$\geq 70\%$ Score of 2	$\geq 60\%$ Score of 1	$< 60\%$ Score of 0
4. Districts have a low proportion of over-identification of teachers for designation (REM).		$\leq 10\%$ Score of 3	$\leq 20\%$ Score of 2	$\leq 30\%$ Score of 1	$> 30\%$ Score of 0

C. Degree of reliability for observation and growth judgements

This check is intended to confirm that observation ratings and student performance are determined in a consistent manner across campus and teaching assignment.¹

Most evidence supports the accuracy of judgements	Some evidence points toward the accuracy of judgements	Limited evidence supports the accuracy of judgements		None or almost no evidence supports judgements
Score of 3	Score of 2	Score of 1		Score of 0
5. <u>Across campuses</u> , observation scores are similar for teachers in REM groups.	$p. \omega^2 \leq 0.01$	$p. \omega^2 \leq 0.06$	$p. \omega^2 \leq 0.14$	$p. \omega^2 > 0.14$
	Score of 3	Score of 2	Score of 1	Score of 0
6. <u>Across campuses</u> , percentages of student growth are similar for teachers in REM groups.	$p. \omega^2 \leq 0.01$	$p. \omega^2 \leq 0.06$	$p. \omega^2 \leq 0.14$	$p. \omega^2 > 0.14$
	Score of 3	Score of 2	Score of 1	Score of 0
7. <u>Across assignments</u> , observation scores are similar for teachers in REM groups.	$p. \omega^2 \leq 0.01$	$p. \omega^2 \leq 0.06$	$p. \omega^2 \leq 0.14$	$p. \omega^2 > 0.14$
	Score of 3	Score of 2	Score of 1	Score of 0
8. <u>Across assignments</u> , percentages of student growth are similar for teachers in REM groups.	$p. \omega^2 \leq 0.01$	$p. \omega^2 \leq 0.06$	$p. \omega^2 \leq 0.14$	$p. \omega^2 > 0.14$
	Score of 3	Score of 2	Score of 1	Score of 0

¹ Observation and growth should be equal when compared across campuses and assignments. A smaller effect-size indicates small differences, thus a greater degree of agreement. A larger effect-size indicates larger differences, thus a smaller degree of agreement.

D. Comparison of district designation percentage to statewide performance standards

This check is intended to confirm that designation rates in each district are aligned with statewide projections of the proportion of designated teachers in each district. For the current year, the checks associated with #11 will be supplemental, to provide information that may be used to refine the designation system.

Most evidence supports the accuracy of judgements	Some evidence points toward the accuracy of judgements	Limited evidence supports the accuracy of judgements	None or almost no evidence supports judgements		
Score of 3	Score of 2	Score of 1	Score of 0		
9. <i>Percentage of students who meet or exceed expected growth</i> in the district is approximately equal to the statewide performance standards for student growth in each of the teacher-designation levels (REM).		≥ 70%	≥ 65%	≥ 60%	< 60%
		Score of 3	Score of 2	Score of 1	Score of 0
10. <i>Observation ratings</i> in the district are approximately equal to the statewide performance standards for teaching proficiency in each of the REM levels.		≥ 80%	≥ 70%	≥ 60%	< 60%
		Score of 3	Score of 2	Score of 1	Score of 0
11. The proportion of teachers on district campuses who are designated as <i>Recognized</i> , <i>Exemplary</i> , or <i>Master</i> is roughly equivalent to other campuses in the same Domain 2A designation group.		$W \leq 0.10$	$w \leq 0.30$	$w \leq 0.50$	$w > 0.50$
		Score of 3	Score of 2	Score of 1	Score of 0

Check 1. The correlation coefficient between observation and growth among all eligible teachers is within the range of expected magnitude reported in research literature.

Pearson product-moment correlation coefficient (r) is calculated between the teacher observation and growth scores of all eligible teachers. Pearson's coefficient is a measure of the strength and direction of linear association between two variables, which can be written as:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

where n is the sample size; x_i and y_i are the person i 's values on x and y (e.g., x = observation score, y = growth score); and \bar{x} and \bar{y} are the sample means of x and y .

Correlation coefficient has a value between -1 (a perfect negative correlation) and $+1$ (a perfect positive correlation). A positive correlation indicates a positive relationship while a negative correlation signifies a negative relationship. For example, when teachers with higher observation scores show higher growth scores, the correlation will be positive; in contrast, when teachers with higher observation scores show lower growth scores, the correlation will be negative. Two correlations with the same numerical value have the same strength whether the correlation is positive or negative. A zero correlation indicates no relationship between the variables. The following guidelines are useful when determining the strength of a correlation: ± 0.1 (small), ± 0.3 (moderate), and ± 0.5 (large) (Cohen, 1988, 1992).

Check 2. The correlation coefficient between district teacher designations and the designation equivalent based on VAM.

Kendall rank correlation coefficient (τ) is calculated between the designation level that the district has made for their teachers (Master, Exemplary, or Recognized) and the same teachers' designation level that is determined by their value-added (VAM) score in SY2020-2021 (Master, Exemplary, Recognized, or Not Designated). Kendall's coefficient is a measure of the strength and direction of ordinal association between two variables, which can be written as:

$$\tau_{xy} = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}$$

where n is the sample size; $n_0 = n(n-1)/2$; $n_1 = \sum_i t_i(t_i-1)/2$; $n_2 = \sum_j u_j(u_j-1)/2$; n_c is the number of concordant pairs; n_d is the number of discordant pairs; t_i is the number of tied values in the i^{th} group of ties for the first quantity; and u_j is the number of tied values in the j^{th} group of ties for the second quantity. Any pair of observations (x_i, y_i) and (x_j, y_j) , where $i < j$, are said to be concordant if the sort of (x_i, y_i) and (x_j, y_j) agrees—that is, if either both $x_i > x_j$ and $y_i > y_j$ holds or both $x_i < x_j$ and $y_i < y_j$. Otherwise, they are said to be discordant.

For example, the correlation will be $+1$ (a perfect positive correlation) when the agreement between the district's designation and VAM-based designation is perfect (i.e., the two rankings are the same). The correlation will be positive when the two designations are similar. The correlation will be -1 (a perfect

negative correlation) when the disagreement between the two designations is perfect (i.e., one ranking is the reserve of the other). When the two designations are independent, then the correlation will be approximately zero.

Check 3. Districts accurately identify teachers for designation based on statewide VAM (REM).

Check 4. Districts have a low proportion of over-identification of teachers for designation (REM).

For teachers who earned a designation in the district (Master, Exemplary, or Recognized), it is determined whether their VAM score (2020-2021) is in the same, higher, or lower designation category based on the state published cut-points for those categories. “Accurate” identification is defined as the VAM score being at or above the state published cut-point of the district designation category, while “over-”identification is defined as the VAM score being below the cut-point. The percentages of “accurate” identification (Check 3) and “over-”identification (Check 4) are calculated.

Check 5. Across campuses, observation scores are similar for teachers in REM groups.

Check 6. Across campuses, percentages of student growth are similar for teachers in REM groups.

Analysis of variance (ANOVA) is performed to compare teachers’ observation score (Check 5) or growth score (Check 6) across different campuses. The analysis model includes the main effects of campus and teacher designation (Master, Exemplary, or Recognized) as well as their interaction effect. Then, partial omega-squared (ω^2) for the campus effect is calculated. Partial omega-squared is a measure of standardized group difference (effect size)—the proportion of the variance in a dependent variable (e.g., observation or growth score) that is accounted for by the independent variable (e.g., campus), with other effects (terms) in the model partialled out of both the dependent variable and the independent variable. It can be written as:

$$\text{partial } \omega^2 = \frac{df_{\text{effect}} (MS_{\text{effect}} - MS_{\text{error}})}{df_{\text{effect}} MS_{\text{effect}} + (N - df_{\text{effect}}) MS_{\text{error}}}$$

where N is the sample size; df is the degrees of freedom; MS_{effect} is the mean sum of squares for the independent variable; and MS_{error} is the mean sum of squares for the error. (Partial) omega-squared is widely viewed as a lesser biased alternative to (partial) eta-squared, especially when sample sizes are small.

Partial omega-squared can have a value between -1 and $+1$. The following guidelines are useful when determining the strength of a partial omega-squared: 0.01 (small), 0.06 (moderate), and 0.14 (large) (Cohen, 1988, 1992). A zero or negative value indicates no effect of the independent variable when controlling for the other effects included in the model.

Check 7. Across assignments, observation scores are similar for teachers in REM groups.

Check 8. Across assignments, percentages of student growth are similar for teachers in REM groups.

ANOVA is performed to compare teachers’ observation score (Check 7) or growth score (Check 8) across different teaching assignments. Teaching assignment is defined as eligible teacher group (possibly, 2-10 groups); or defined as tested subjects, non-tested subjects, or both subjects when there is only one eligible teacher group. The analysis model includes the main effects of teaching assignment and teacher designation (Master, Exemplary, or Recognized) as well as their interaction effect. Then, partial omega-squared (ω^2) for the teaching assignment effect is calculated.

Check 9. Percentage of students who meet or exceed expected growth in the district is approximately equal to the statewide performance standards for student growth in each of the teacher-designation levels (REM).

Check 10. Observation ratings in the district are approximately equal to the statewide performance standards for teaching proficiency in each of the REM levels.

For teachers who earned a designation in the district (Master, Exemplary, or Recognized), it is determined how close their growth score (Check 9) or observation score (Check 10) is to the published cut-point that corresponds to their designation category. A “closeness” score based on the proximity of the growth score or observation score to the corresponding performance standard at each designation level is established on a 0-100% scale. The score value is calculated using an exponential equation that assigns a score based on the proximity of each teacher’s score to the corresponding performance standard. More points are given when the score is closer to the performance standard. After a score has been assigned to each teacher, these scores are averaged within each of the REM levels. Finally, an overall mean value is calculated based on the averages within the designation groups. The state published cut-points used are shown below:

<i>Growth standard group</i>	<i>% of students meeting or exceeding growth targets</i>
Recognized	55%
Exemplary	60%
Master	70%

<i>Observation standard group</i>	<i>Based on T-TESS</i>	<i>Based on another rubric</i>
Recognized	3.7	74% of points
Exemplary	3.9	78% of points
Master	4.5	90% of points

Check 11. The proportion of teachers on district campuses who are designated as Recognized, Exemplary, or Master is roughly equivalent to other campuses in the same Domain 2A designation group.

The district percentage of (1) Master designation, (2) Exemplary or higher designations, or (3) Recognized or higher designations in 2018-2019 are each compared to a State average of campuses within each of the Domain 2A categories. In other words, the district and state percentages are obtained for teachers within Domain 2A A-rated campuses, Domain 2A B-rated campus, etc. Then, Cohen’s w is calculated from each possible comparison in the Domain 2A categories and a weighted mean value is calculated over the (1), (2), and (3) designation levels.

Cohen’s w is a measure of association between two nominal variables. With a binary outcome (e.g., designated vs. not designated), it can be written as follows with directionality considered:

$$w = \text{sign}(p_1 - p_0) \sqrt{\frac{(p_1 - p_0)^2}{p_0} + \frac{(p_1 - p_0)^2}{(1 - p_0)}}$$

where p_1 is the district percentage and p_0 is the statewide expected percentage. The value will be 0 when the district percentage is equal to the statewide percentage for a Domain 2A category. In

contrast, the value will be positive when the district percentage is larger than the statewide percentage; or it will be set to zero when the district percentage is smaller than the statewide percentage.