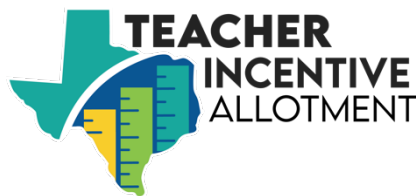**Summary of the Step 2 Data Validation Process**

Cohort A, B, C & D

Year 4 – 2022 Data Submission

Table of Contents

Greetings, LEA Colleagues,

Thank you for participating in the Teacher Incentive Allotment. We recognize the significance of this decision and admire your willingness to benefit students, teachers, and the LEA through this program.

Purpose of this document

This document provides a *summary* of the Step 2 Data Validation process in three sections.

- The first is a reader-friendly description of the statistics used in each check and the type of evidence each check provides toward validating a district's system for designation.
- The second section contains the Validation Scoring Rubric. This rubric is used to summarize evidence and assign points for each check. Ultimately, scores on the rubric allow the Texas Education Agency (TEA) to determine if sufficient evidence exists to support the conclusion that a district's designation system will result in valid teacher selections. A district need not score perfectly on every dimension, but component scores, taken together, help TEA decide about a district's system.
- For those interested in a deeper understanding of the data 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.

This resource is intended to provide a big picture perspective of the checks performed by Texas Tech on the data submitted by LEAs. We recommend that you read this document first before reviewing the data template, and keeping it handy while your data file is being prepared. If you have questions about this summary, please contact TIA@ttu.edu.

A procedural change

If your district previously participated in TIA or joined trainings related to Step 2 Data Validation, you may notice some changes to the checks. These changes are outlined in an appendix to this letter on the following page. In general, one new check was created (#3), a new supplemental check was added (#11), and the weighting of the checks was revised.

Finally, one other change was made to help eliminate reporting errors.

- *Converting observation scores.* This change pertains **only** to districts with an observation rubric having a minimum indicator score of 0 (zero). **Note, for districts using T-TESS, this change will not affect your data submission process.** In the past, we requested that such districts increase all indicator scores by 1 (one) to facilitate calculation of check scores. This year, however, we ask that all scores be submitted as gathered by the district, and Texas Tech will convert them as necessary to a scale beginning at 1.

If you have questions about this summary or these changes, please contact TIA@ttu.edu.

Sincerely,

Texas Tech Team for TIA

iv

Summary of Changes

Changes to Data Validation Checks (2022)

Overview of Changes:

Current Check Number	Domain	New Check Number	Weight	Change
1	A	1	6	No change
2	B	2	6	Changed weight
3	B	Removed	NA	Removed
4	B	Removed	NA	Removed
NA	B	3	4	New Check 3 using proximity to designation if determined by statewide VAM-based on point matrix, additional details below
5	C	4	2	Changed numbering
6	C	5	2	Changed numbering
7	C	6	2	Changed numbering
8	C	7	2	Changed numbering
9	D	8	1	Changed numbering
10	D	9	1	Changed numbering
11 (supp)	E	10 (supp)	0	Changed numbering
NA	E	11 (supp)	0	New Supplemental Observation check

New Check 3:

For each designated teacher, we will calculate the proximity of district designation to the same teacher's designation as determined by the state-wide Value Added Model (VAM). Scores for individual teachers will be assigned as described in the table below. Then all scores will be averaged to assign rubric points. For example, if a district designates a teacher as *Exemplary,* while the state-wide VAM determines the teacher is *Recognized*, the district would receive a score of .75 for that teacher. All calculated teachers scores would be averaged together for the final score for Check 3, which would then be converted based on the cut points below.

District Designations	Designations if determined by the state-wide VAM			
	Not Designated	*Recognized*	*Exemplary*	*Master*
Recognized	0.00	**1.00**	0.75	0.50
Exemplary	-0.25	0.75	**1.00**	0.75
Master	-1.00	0.25	.75	**1.00**

Points will be assigned to results from the average based on these cut points:

≥ .70	≥ .30	> 0	≤ 0
Score of 3	**Score of 2**	**Score of 1**	**Score of 0**

New Supplemental (unscored) check 11:

This will be the percent of max observation score Standard Deviation with the following cut points. This check is intended to help system administrators gauge the extent to which observation and appraisal practices distinguish instruction that is more or less effective. For this year, this check is unscored and may be considered as a scored check in the future.

$\sigma \geq 0.12$	$\sigma \geq 0.10$	$\sigma \geq 0.08$	$\sigma < 0.08$
Score of 3	**Score of 2**	**Score of 1**	**Score of 0**

Reader-Friendly Description of Validation Checks

The analyses described below are intended to validate 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 confirm or reject designation of individual teachers. Meeting these standards allows stakeholders to have confidence that the designation system is fair and accurate.

Domain A. Correlation between teacher observation ratings and student performance ratings
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.
For this check, analysts calculate the correlation coefficient (Pearson 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 gain 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 provide one piece of evidence about the validity of the designation system.

Domain B. Confirm relation between district designations and student growth calculations
Check 2 District designations of Recognized, Exemplary, and Master (REM) teachers are found in similar proportion to designations as determined by the state-wide VAM.
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 district data from SY2021-2022 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 designation as determined by the state-wide VAM than those with an Exemplary and that teachers with this designation would have a higher-ranking designation as determined by the state-wide VAM than those with a Recognized designation. Results from this analysis provide another piece of evidence about the validity of the designation system.
Check 3 District designation decisions for REM teachers, in tested subjects, are in proximity to designations as determined by the state-wide VAM.
For this check, analysts examine the accuracy with which local designation systems designate each eligible teacher in a tested subject based on calculations of the designations if they were determined by the state-wide VAM. The table below shows how scores are calculated based on the proximity of district designation to the designation if it was determined by the state-wide VAM. Scores reflect a positive value for accurate designation, and a negative value for decisions that are not aligned with designations if they had been determined by the state-wide VAM. For example, if a local system designates a teacher as Exemplary, and this designation is concurrent with the designation determined by the state-wide VAM, then an accuracy score of 1.00 is assigned. On the other hand, if a local system designates a teacher as Master, but the designation as determined by the state-wide VAM indicates no designation should be made, then an accuracy score of -1.00 is assigned. The expectation is that local systems will accurately identify teachers, and their levels, for designation. This analysis provides evidence about the concurrent validity of the local designation system.

	Designations if determined by the state-wide VAM			
District Designations	*Not Designated*	*Recognized*	*Exemplary*	*Master*
Recognized	0.00	**1.00**	0.75	0.50
Exemplary	-0.25	0.75	**1.00**	0.75
Master	-1.00	0.25	.75	**1.00**

Domain C. Degree of reliability for observation and growth judgements
Check 4 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. 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 5 Across campuses, percentages of student growth are similar for teachers in REM groups.
This check is like Check 4 in that 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; 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 6 Across assignments, observation scores are similar for teachers in REM groups.
This check is also like Check 4, but instead of making comparisons across campuses, it looks for similarities in observation ratings within REM groups across teaching assignment. As before, 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 the data 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 7 Across assignments, percentages of student growth are similar for teachers in REM groups.
This check is like Check 5, but instead of making comparisons in observation, the comparison is of student growth (percentage of students who meet or exceed learning expectations) within the REM 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.

Domain D. Comparison of district designation percentage to statewide performance standards
Check 8 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 9** Observation ratings in the district are approximately equal to the statewide performance standards for teaching proficiency in each of the teacher-designation levels (REM).
Both checks involve simple comparisons between statewide performance standards for each designation level and district-level results.   Performance standards were 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 mismatch. 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 mismatch. 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.

Domain E. Supplemental System Checks (not scored)
Check 10 The proportion of teachers on district campuses who are designated as Recognized, Exemplary, or Master is roughly equivalent to other campuses with the same Domain 2A rating.
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 like 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. For the 2022 Step 2 Data Validation process, TEA will use 2021-22 Domain 2a ratings. Any district submitting a campus receiving a "Not Rated" label will be exclude from this analysis for comparison purposes. 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.
Check 11 The variability in observation ratings among all eligible teachers is within the range of historical magnitude.
The purpose of this check is to display for leaders of the district's TIA initiative the spread of teacher observation scores gathered during the data capture year. For this check, analysts calculate standard deviation of (max-scaled) observation scores among all eligible teachers. The expectation is that observation scores for the district will be distributed in a manner that gives some evidence about the ability of the local system to differentiate between ineffective and effective instruction.

Validation Rubric for TIA Step 2 Data Validation

Domain 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	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**
1. The correlation coefficient between observation and growth among all *eligible* teachers is within the range of expected magnitude reported in research literature. *Earned points x 6 = weighted score for this check*	$r \geq .24$ **Score of 3** $r \geq .16$ **Score of 2**	$r \geq .08$ **Score of 1**	$r < .08$ **Score of 0**

Domain 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 SY2021-2022 VAM data.

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**
2. District designations of REM teachers are found in similar proportion to designations as determined by the state-wide VAM. *Earned points x 6 = weighted score for this check*	$\tau \geq 0.50$ **Score of 3** $\tau \geq 0.30$ **Score of 2**	$\tau \geq 0.10$ **Score of 1**	$\tau < 0.10$ **Score of 0**
3. District designations of REM teachers, in tested subjects, are in proximity to designations as determined by the state-wide VAM. *Earned points x 4 = weighted score for this check*	$\geq .70$ **Score of 3** $\geq .30$ **Score of 2**	> 0 **Score of 1**	≤ 0 **Score of 0**

Domain C. Degree of reliability for observation and growth judgements

These checks are 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
4. <u>Across campuses</u>, observation scores are similar for teachers in REM groups. *Earned points x 2 = weighted score for this check*	sp. $\omega^2 \le 0.01$ **Score of 3** — sp. $\omega^2 \le 0.06$ **Score of 2**	sp. $\omega^2 \le 0.14$ **Score of 1**	sp. $\omega^2 > 0.14$ **Score of 0**
5. <u>Across campuses</u>, percentages of student growth are similar for teachers in REM groups. *Earned points x 2 = weighted score for this check*	sp. $\omega^2 \le 0.01$ **Score of 3** — sp. $\omega^2 \le 0.06$ **Score of 2**	sp. $\omega^2 \le 0.14$ **Score of 1**	sp. $\omega^2 > 0.14$ **Score of 0**
6. <u>Across assignments</u>, observation scores are similar for teachers in REM groups. *Earned points x 2 = weighted score for this check*	sp. $\omega^2 \le 0.01$ **Score of 3** — sp. $\omega^2 \le 0.06$ **Score of 2**	sp. $\omega^2 \le 0.14$ **Score of 1**	sp. $\omega^2 > 0.14$ **Score of 0**
7. <u>Across assignments</u>, percentages of student growth are similar for teachers in REM groups. *Earned points x 2 = weighted score for this check*	sp. $\omega^2 \le 0.01$ **Score of 3** — sp. $\omega^2 \le 0.06$ **Score of 2**	sp. $\omega^2 \le 0.14$ **Score of 1**	sp. $\omega^2 > 0.14$ **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 level of agreement. A larger effect-size indicates larger differences, thus a smaller level of agreement.

Domain D. Comparison of district designation percentage to statewide performance standards

These checks are intended to confirm that designation rates in each district are aligned with statewide projections of the proportion of designated teachers in each district.

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

	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
8. *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). *Earned points x 2 = weighted score for this check*	≥ 70% **Score of 3**	≥ 65% **Score of 2**	≥ 60% **Score of 1**	< 60% **Score of 0**
9. *Observation ratings* in the district are approximately equal to the statewide performance standards for teaching proficiency in each of the REM levels. *Earned points x 1 = weighted score for this check*	≥ 80% **Score of 3**	≥ 70% **Score of 2**	≥ 60% **Score of 1**	< 60% **Score of 0**

Domain E. Supplemental Checks

These checks are intended to provide additional, non-scored evidence to districts about the validity of their local designation system. Check #10 reflects the degree to which designation decisions are comparable among districts with the same Domain 2A ratings. Check #11 shows the variance in district's teacher observation scores as an indicator of the extent to which observers differentiate between more effective and less effective instruction. For the current year, these checks are supplemental and are not factored into data validation scores or system validation decisions

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

	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
10. The proportion of teachers on district campuses who are designated as *Recognized*, *Exemplary*, or *Master* is roughly equivalent to other campuses with the same Domain 2A rating. *No points assigned for supplemental check.*	$w \leq 0.10$ **Score of 3**	$w \leq 0.30$ **Score of 2**	$w \leq 0.50$ **Score of 1**	$w > 0.50$ **Score of 0**
11. The variability in observation ratings among all eligible teachers is within the range of expected magnitude. *No points assigned for supplemental check.*	$\sigma \geq 0.12$ **Score of 3**	$\sigma \geq 0.10$ **Score of 2**	$\sigma \geq 0.08$ **Score of 1**	$\sigma < 0.08$ **Score of 0**

Description of Statistical Analysis Protocols

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. **District designations of REM teachers are found in similar proportion to designations as determined by the state-wide 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 (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; $n_0 = \frac{n(n-1)}{2}$; $n_1 = \sum_i \frac{t_i(t_i - 1)}{2}$; $n_2 = \sum_j \frac{u_j(u_j - 1)}{2}$; 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 designations if determined by the state-wide VAM 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 reverse of the other). When the two designations are independent, then the correlation will be approximately zero.

Check 3. *District designation decisions for REM teachers, in tested subjects, are in proximity to designations as determined by the state-wide VAM.*

For teachers of tested subjects who earned a designation (Master, Exemplary, or Recognized) in the district, it is determined whether the district designation is the same, higher, or lower than the designation if it were determined by the state-wide VAM. An "accuracy" score ranging from -1.00 to +1.00 is assigned based on the proximity between the district designation and the designation if it were determined by the state-wide VAM. The table below shows how values are assigned based on proximity:

	Designations if determined by the statewide VAM			
District Designations	*Not Designated*	*Recognized*	*Exemplary*	*Master*
Recognized	0.00	**1.00**	0.75	0.50
Exemplary	-0.25	0.75	**1.00**	0.75
Master	-1.00	0.25	.75	**1.00**

More points are given when the district designation is closer to the designations if determined by the state-wide VAM. After a score has been assigned to each teacher, these scores are averaged to produce an overall score for the district.

Check 4. *Across campuses, observation scores are similar for teachers in REM groups.*
Check 5. *Across campuses, percentages of student growth are similar for teachers in REM groups.*
ANOVA is performed to compare teachers' observation score (Check 4) or growth score (Check 5) across different campuses. The analysis model includes the main effects of campus and teacher designation (Master, Exemplary, Recognized) as well as their interaction effect. Then, semi partial omega-squared (ω^2) for the campus effect is calculated. Semi 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 partialed out of the independent variable. It can be written as:

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

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. (Semi partial) omega-squared is widely viewed as a lesser biased alternative to (semi partial) eta-squared, especially when sample sizes are small.

Semi partial omega-squared can have a value between –1 and +1. The following guidelines are useful when determining the strength of a semi 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 6. *Across assignments, observation scores are similar for teachers in REM groups.*
Check 7. *Across assignments, percentages of student growth are similar for teachers in REM groups.*
ANOVA is performed to compare teachers' observation score (Check 6) or growth score (Check 7) across different teaching assignments. Teaching assignment is defined as two or more eligible teacher 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 8**. 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 9**. 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 8) or observation score (Check 9) 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:

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

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

The exponential equations used are shown below:

In Check 8
For Master teachers,

$$s_i = f(x_i) + g(x_i)\left(\frac{x_i-0.5}{0.7-0.5}\right)^2$$

$$f(x_i) = \begin{cases} 1 & 0.7 \le x_i \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.5 \le x_i < 0.7 \\ 0 & \text{otherwise} \end{cases};$$

For Exemplary teachers,

$$s_i = f(x_i)\left(\frac{x_i-0.5}{0.6-0.5}\right)^2 + g(x_i) + h(x_i)\left(1 - \frac{x_i-0.7}{1-0.7}\right)^2,$$

$$f(x_i) = \begin{cases} 1 & 0.5 \le x_i < 0.6 \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.6 \le x_i < 0.7 \\ 0 & \text{otherwise} \end{cases}, h(x_i) = \begin{cases} 1 & 0.7 \le x_i \\ 0 & \text{otherwise} \end{cases};$$

For Recognized teachers,

$$s_i = f(x_i)\left(\frac{x_i-0.5}{0.55-0.5}\right)^2 + g(x_i) + h(x_i)\left(1 - \frac{x_i-0.6}{1-0.6}\right)^2,$$

$$f(x_i) = \begin{cases} 1 & 0.5 \le x_i < 0.55 \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.55 \le x_i < 0.6 \\ 0 & \text{otherwise} \end{cases}, h(x_i) = \begin{cases} 1 & 0.6 \le x_i \\ 0 & \text{otherwise} \end{cases};$$

where s_i and x_i are the person i's values on closeness score and growth score, respectively.

In Check 9
For Master teachers,

$$s_i = f(x_i) + g(x_i)\left(\frac{x_i-0.7}{0.9-0.7}\right)^2,$$

$$f(x_i) = \begin{cases} 1 & 0.9 \le x_i \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.7 \le x_i < 0.9 \\ 0 & \text{otherwise} \end{cases};$$

For Exemplary teachers,

$$s_i = f(x_i)\left(\frac{x_i-0.7}{0.78-0.7}\right)^2 + g(x_i) + h(x_i)\left(1 - \frac{x_i-0.9}{1-0.9}\right)^2,$$

$$f(x_i) = \begin{cases} 1 & 0.7 \le x_i < 0.78 \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.78 \le x_i < 0.9 \\ 0 & \text{otherwise} \end{cases}, h(x_i) = \begin{cases} 1 & 0.9 \le x_i \\ 0 & \text{otherwise} \end{cases};$$

For Recognized teachers,

$$s_i = f(x_i)\left(\frac{x_i-0.7}{0.74-0.7}\right)^2 + g(x_i) + h(x_i)\left(1 - \frac{x_i-0.78}{1-0.78}\right)^2,$$

$$f(x_i) = \begin{cases} 1 & 0.7 \le x_i < 0.74 \\ 0 & \text{otherwise} \end{cases}, g(x_i) = \begin{cases} 1 & 0.74 \le x_i < 0.78 \\ 0 & \text{otherwise} \end{cases}, h(x_i) = \begin{cases} 1 & 0.78 \le x_i \\ 0 & \text{otherwise} \end{cases};$$

where s_i and x_i are the person i's values on closeness score and observation score, respectively.

Check 10. 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 rating.
The campus cumulative percentage of (1) Master designation, (2) Exemplary or higher designations, or (3) Recognized or higher designations 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 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.

Check 11. The variability in observation ratings among all eligible teachers is within the range of expected magnitude.
Standard deviation (σ) is calculated for the (max-scaled) observation score of all eligible teachers. Standard deviation is a measure of variation or dispersion of a variable, which can be written as:

$$\sigma = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

where *n* is the sample size; x_i is the person *i*'s values on *x* (e.g., observation score); and $\bar{x}$ is the sample mean of *x*. A low standard deviation indicates that teachers' observation scores are close each other and to the mean, while a high standard deviation indicates that scores are spread out over a wider range.