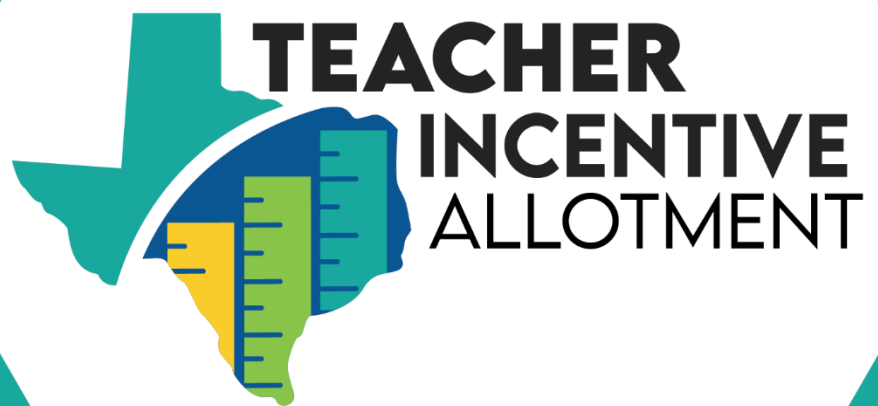*Killeen ISD*

STEP 2 Data Validation Report

Spring 2023

For questions about the analysis and results,

contact TIA@ttu.edu.

Purpose and Process of Data Validation

District approval to designate teachers involves a two-step review process. The first step is carried out by the Texas Education Agency (TEA). The second step is performed by TEA in collaboration with Texas Tech University. TEA reviews the results from Step 1 and Step 2 to determine final approval.

This report pertains to the second step in the approval process. In Step 2, data provided by districts about designated and all designation-eligible teachers are examined to verify that the district system produced reliable results, and that there is a consistent, positive association between teacher observation scores and measures of student growth. Additional analyses comparing district results to state-level results are performed to ensure that teachers earning district designations are indeed among the most effective in the State. Based on feedback from previous cohorts, we have added some general guidance and suggestions for next steps after leaders have reviewed the scoring for the district's local designation system (see Appendix A).

Verifying District Results

In Year 4, TEA and Texas Tech operationalized the method for verification by using 11 indicators as checks for evidence of the reliability and validity of district systems. These 11 checks were compiled into a rubric that Texas Tech uses to evaluate district data (see Appendix B). There are five main domains for these checks that include examining:

A. The association between observation and district-reported student growth.
B. The relation of district-reported student growth and statewide value-added scores.
C. The extent to which observation and district-reported student growth are equivalent for teachers in designation categories across campuses and teaching assignments.
D. The extent to which district designation patterns conform to patterns found in state-level analyses of growth and observation.
E. Supplemental checks that examine similarity of designation patterns among districts with similarly Domain 2A ratings, and a visual depiction of the relation between the distribution of observation and growth scores.

Scores on the supplemental checks do not count towards or take away from the summative Step 2 score. Instead, the supplemental checks are intended to support future efforts to improve district local designation systems. Note, in the future years, the 11 checks are subject to change.

Once analysts have completed the 11 checks, scores are assigned in the rubric. Some scores are weighted more heavily than others. Greater weights reflect the relative value of each check as it pertains to the intended functioning of the designation system. The rubric enables analysts to produce a summative score for the district, and provides data needed to inform a judgment about whether the district system does indeed result in the identification and designation of the most effective teachers in the State. Additional technical information is provided in Appendix C about the specific statistical analysis performed for each check.

Results for Killeen ISD

Descriptive Statistics from Data Submission

Basic information about Killeen ISD and the data submitted for verification are summarized in Table 1. This table contains information including district name (and number) along with the date on which data files were transferred to Texas Tech University for evaluation, the name of the rubric used for teacher observation, and the number of designated and teachers in eligible teaching assignments included in the files.

Table 1.
Basic Information of Data Submission

District Name:	**Killeen ISD**
District Number:	**14906**
Number of Errors in Data Submission:	**957**
Number of Attempted Submissions:	**4**
Date of Initial File Transfer:	**October 17, 2022**
Date of Final File Transfer:	**October 17, 2022**
Observation Rubric Used:	**T-TESS**
Number of Designated Teachers:	**306**
Number of Designation-Eligible Teachers:	**1340**

Table 2 also contains district information about the number and the percentage of teachers designated by level. Of the 1340 teachers in eligible teaching assignment, 306 (23 %) were identified for designation.

Table 2.
Frequency of Teacher Designation

Category	*n*	*%*
Master	24	2 %
Exemplary	122	9 %
Recognized	160	12 %
Non-designated	1034	77 %
Total	1340	100 %

Data Submission Errors

Killeen ISD had 957 total errors, passed all data checks in 4 round(s), and had 1 type(s) of error(s) during the data submission process. Data submission through the TTU Data Portal and TEA's alignment review ensured that a technically accurate file was put forth and that a district implemented their system with fidelity according to their system application. The results produced during the data submission review did not affect the scores during data validation. Note, the following errors occurred (error category and count): Main Grade Level(s) for Observation and Growth error (957 instances). Please ensure the submission of accurate data for future years.

Data Validation Scores

Table 3 contains results from all the verification checks along with the rubric score and the resulting weighted score. For your district, the total point value possible was 78. When district data sets were limited and the analysis associated with the check could not be performed, the check and associated points were omitted. The omission of checks does not negatively impact the district's score. The notation of N/A in Table 3, if present, indicates omitted checks.

The total verification score for Killeen ISD was 48 out of 78 possible points, or 61 %. Based on holistic review of your application system and the results in this report, this is a passing score for the data generated by the district's designation system.

Table 3.
Step Two Verification Scores

Domain	Check	Possible Points	Results	Score	Weight	Score × Weight
A. Correlation between teacher observation ratings and student growth ratings	C1	0-3	$r = 0.34$	3	× 6	18
B. Relationship between teacher designations and VAM	C2	0-3	$\tau = 0.21$	1	× 6	6
	C3	0-3	0.16	1	× 4	4
C. Degree of reliability for observation and growth judgements	C4	0-3	sp. $\omega^2 = 0.06$	1	× 2	2
	C5	0-3	sp. $\omega^2 = -0.02$	3	× 2	6
	C6	0-3	sp. $\omega^2 = -0.01$	3	× 2	6
	C7	0-3	sp. $\omega^2 = 0.20$	0	× 2	0
D. Comparison of district designation percentage to statewide performance standards	C8	0-3	79 %	3	× 1	3
	C9	0-3	87 %	3	× 1	3
E. Supplemental checks	C10	0-3	$w = 0.00$	3	× 0	–
	C11	0-3	$\sigma = 0.12$	3	× 0	–
					Total	**48 / 78**

Note. r = Pearson product-moment correlation coefficient; τ = Kendall rank correlation coefficient; sp. ω^2 = semi-partial omega-squared; σ = standard deviation.

Distribution of and Relation Between Teacher Observation and Student Growth

The figures below provide a representation of how teacher observation ratings and student growth are distributed in the district's data submission of designation-eligible teachers this year. Please note, teacher observation ratings have been converted to the proportion of total points earned on the district's appraisal rubric. Figure 1 can be examined to determine the extent to which the distribution of observation ratings (blue bars) follow a pattern similar to the distribution of student growth percentages (red bars). Figure 2 provides a 1:1 match between growth and observation and can be examined to determine the extent to which teachers who have higher percentages of students meeting or exceeding growth targets also receive higher appraisal ratings.

Figure 1. *Distribution of Percentage of Students Meeting Growth Standards and Teacher Observation Ratings*
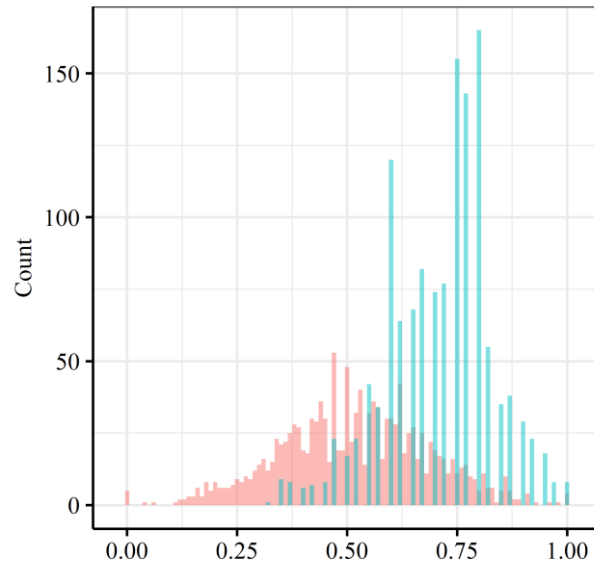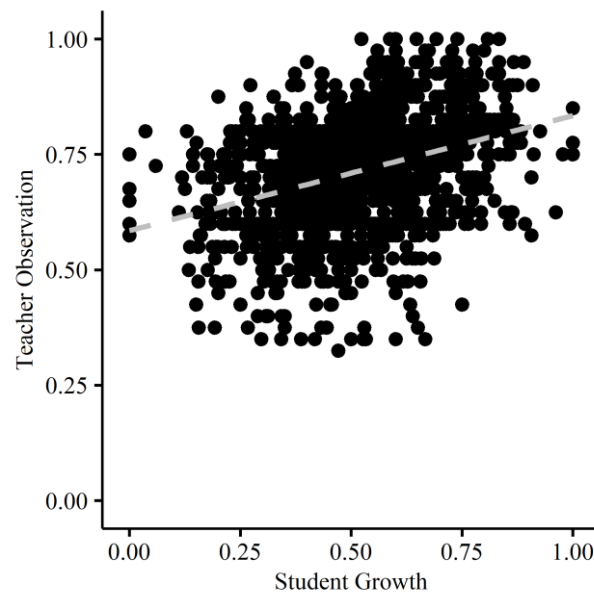

Figure 2. *A 1:1 Match of Individual Teachers' Observation and Growth*

Appendix A

Analyzing LEA Rubric Scores: Considerations for Next Steps

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.

Next Steps:
- Dive deeper into your districts correlation to uncover areas of focus to improve on. This could be a particular campus/appraiser, student growth measure, TIA teacher category or even a subject/grade-level
- Look for areas of inflation in your teacher observation and growth scores.

Questions to Consider:
- Do flaws exist in your appraisal system that require recalibration or more training?
- Do flaws exist in the construction, administration and/or calculation of your student growth measure?
- Is inflation in teacher observation and growth scores happening at a particular campus, a specific subject/grade level, or maybe with a certain appraiser.

Resources:
- Excel Analysis Tool
- Correlation Training
- Correlation Resources

Keep in mind that the root of the issue could be teacher observation, student growth or both.

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.

Next Steps:
- Dive deeper into your district's distribution of both observation and student growth scores.
- Dive deeper into the total overall count of designations and by designation level.

Questions to Consider:
- Is there enough stretch in the data to accurately distinguish between effective and ineffective teachers?
- What other components are contributing to designation decisions? Is this causing problems in my designation system?
- Are the same students being looked at for my growth measure as STAAR?
- Are you measuring growth in a consistent way across all growth measures?
- Are you administering the growth measure with the same fidelity as STAAR?
- How do the number of designations you issued compared to the state? Does this percent of designations make sense for my district compared to my performance in the state?

Resources:
- SAS EVAAS Manual
- Walkthrough on the statewide VAM model
- TEA 2022 Accountability Rating System

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.

Next Steps
- Dive deeper into your district's observation and growth scores looking at each designation level as a group.

Questions to Consider:
- Do designated teachers in each respective designation level across campuses, assignments, and teacher categories have similar student growth scores and teacher observation ratings? Do you believe that designation decisions are "fair" for each designation level?
- Do you notice trends at the campus level, teaching assignment level, or teacher category level that are causing designations to appear unfair?
- What type of support would an appraiser need to ensure they are rating teachers fairly?
- Were any trends noted caused because a particular appraiser gave higher scores on a campus relative to other appraisers on that campus or relative to scores on like campuses?
- Was the fidelity in administration of growth measure in one teaching assignment different than in another teaching assignment?
- How consistent was growth measured across all growth measures?
- Are your student growth and teacher observation from the same grade and subject?

Resources:
- Teacher Calibration Protocols
- Teacher Observations for the Teacher Incentive Allotment
- Inter-rater Reliability Webinar
- Inter-Rater Reliability FAQ
- Setting expected growth targets training
- Commonly used student growth assessments

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.

Next Steps:
- Dive deeper into each individual **<u>designated</u>** teacher's observation and growth scores.

Questions to Consider:
- How close are my designated teachers to the statewide performance standards? Are the differences explicable? (e.g., my district is a high performing district compared to that of the average district in the state?)
- Are there other contributing factors in making designation decisions that are causing my designated teachers to have large differences to the statewide performance standards?
- Are the differences in the performances caused by inflated growth scores or observation scores? (e.g., large number of teachers having 100% growth on the SLO growth measure, or large number of teaches receiving all 4's and 5's on T-TESS)

Resources:
- Statewide Performance Standards: Teacher Observation
- Statewide Performance Standards: Student Growth

APPENDIX B

Verification Rubric for TIA Step 2

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 score x 6 = weighted score for this check* *Your district's earned score = 3* *Your district's weighted score = 18*	$r \geq 0.24$ **Score of 3**	$r \geq 0.16$ **Score of 2**	$r \geq 0.08$ **Score of 1** $\quad$ $r < 0.08$ **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 Recognized, Exemplary and Master (REM) teachers are found in similar proportion to designations as determined by the state-wide VAM. *Earned score x 6 = weighted score for this check* *Your district's earned score = 1* *Weighted score = 6*	$\tau \geq 0.50$ **Score of 3**	$\tau \geq 0.30$ **Score of 2**	$\tau \geq 0.10$ **Score of 1** $\quad$ $\tau < 0.10$ **Score of 0**
3. District designations of for REM teachers, in tested subjects, are in proximity to designations as determined by the state-wide VAM. *Earned score x 4 = weighted score for this check* *Your district's earned score = 1* *Your district's weighted score = 4*	≥ 0.70 **Score of 3**	≥ 0.30 **Score of 2**	> 0 **Score of 1** $\quad$ ≤ 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**

	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
4. <u>Across campuses</u>, observation scores are similar for teachers in REM groups. *Earned score x 2 = weighted score for this check* *Your district's earned score = 1* *Your district's weighted score = 2*	sp. $\omega^2 \leq$ 0.01 **Score of 3**	sp. $\omega^2 \leq$ 0.06 **Score of 2**	sp. $\omega^2 \leq$ 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 score x 2 = weighted score for this check* *Your district's earned score = 3* *Your district's weighted score = 6*	sp. $\omega^2 \leq$ 0.01 **Score of 3**	sp. $\omega^2 \leq$ 0.06 **Score of 2**	sp. $\omega^2 \leq$ 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 score x 2 = weighted score for this check* *Your district's earned score = 3* *Your district's weighted score = 6*	sp. $\omega^2 \leq$ 0.01 **Score of 3**	sp. $\omega^2 \leq$ 0.06 **Score of 2**	sp. $\omega^2 \leq$ 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 score x 2 = weighted score for this check* *Your district's earned score = 0* *Your district's weighted score = 0*	sp. $\omega^2 \leq$ 0.01 **Score of 3**	sp. $\omega^2 \leq$ 0.06 **Score of 2**	sp. $\omega^2 \leq$ 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 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			
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 score x 1 = weighted score for this check* *Your district's earned score = 3* *Your district's weighted score = 3*		$\geq 70\%$ **Score of 3**	$\geq 65\%$ **Score of 2**	$\geq 60\%$ **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 score x 1 = weighted score for this check* *Your district's earned score = 3* *Your district's weighted score = 3*		$\geq 80\%$ **Score of 3**	$\geq 70\%$ **Score of 2**	$\geq 60\%$ **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 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			
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 score 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 score 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**	

APPENDIX C

Statistical Analysis Protocols for TIA Step 2

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 Recognized, Exemplary and Master (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 model 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 Recognized, Exemplary, and Master 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 in the district (Master, Exemplary, or Recognized), it is determined whether the district designation is in the same, higher, or lower than the designation if it were determined by the state-wide VAM model. 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 model. 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	0.75	**1.00**

More points are given when the district designation is closer to the designations if determined by the state-wide VAM model. 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.***

Analysis of variance (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:

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

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. (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, semi-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.

<u>Check 10</u>. *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.

<u>Check 11.</u> *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:

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

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.