Teacher Effectiveness Evaluation Model 2021-22

This report describes the Teacher Effectiveness Evaluation Model for SY2021-22. Measuring teacher effectiveness requires multiple measures, both quantitative and qualitative to capture the range of instructional skills used in teaching and to determine how much students benefit academically from their teachers.

The SY2021-22 model is made up of four components including the Danielson Framework, Academic Growth, the Student Survey, and the Teacher Reflection. Each component factors into a teacher's final score, albeit with different weighting. The Danielson Framework comprises the majority of the score determination by making up 67% of the total score. The Academic Growth makes up 20% of the total score. The Student Survey makes up 10% of the total score and the Teacher Reflection is 3% of the total score. Each component is described below and how the points are determined.

Danielson Framework

The Danielson teacher evaluation framework uses 22 criteria nested within four domains. They are: Planning and preparation (N=6); the classroom environment (N=5); instruction (N=5); and professional responsibilities (N=6). Each of the 22 components is scored on a four-point rubric:

1 = Unsatisfactory
2 = Basic
3 = Proficient
4 = Distinguished

The maximum number of points possible on the Danielson is 88 points (22 components X 4 pt. rubric).

Academic Growth

Academic growth will be determined by calculating the growth of TUSD pre-post scores in English Language Arts (ELA), and Math for grades 3-11 from the fall results to the spring results. This approach, however, has some limitations in that the TUSD pre-post-tests in ELA and Math can measure the academic impact of only about a quarter of our teachers (called 'A' teachers). The non-ELA and non-Math teachers (called 'B' teachers) make up the other three-quarters of the teaching core.

   A. What is an 'A' or a 'B' teacher in TUSD?

- An 'A' teacher is any K-2 teacher with fall and spring DIBELS or EDL scores. Grades K-2 will use the DIBELs, EDL or some other assessment to compare the fall results to the spring results.

- An 'A' teacher is also any teacher who teaches math or ELA in grades 3 – 11. Elementary teachers in grades 3 – 5 are 'A' teachers because they teach both math and ELA. All math and ELA teachers in grades 6 – 11 are 'A' teachers.
    - ELA: Grades 3 – 11 are used because they are assessed by TUSD pre-post-tests.
    - Math: Grades 3 – 11 are used because they are assessed by TUSD pre-post-tests.

Teacher Effectiveness Evaluation Model 2021-22

- A 'B' teacher is any teacher who is not an 'A' teacher. For example, if you are a 6th grade science teacher, you are considered a 'B' teacher. If you are a 12th grade AP chemistry teacher, you are also considered a 'B' teacher. The 'B' teachers will be assigned growth points based on the school or the district average.

B. <u>Who will take the assessment:</u> All students in grades K – 2 will take the DIBELs or EDL assessment. In grades, 3 – 11, students will take the TUSD pre-post-test in ELA and math.

C. <u>When will the assessment be administered:</u> DIBELS is typically administered three times a year. The first test in the fall and the last test in the spring will be used. EDL is administered twice a year, once in the fall and once in the spring. TUSD pre-post tests are administered early in the fall and again in mid-spring.

D. <u>Scoring:</u> TUSD pre-post scores used are from the current year so that students are compared against their own scores to measure growth. Teachers who teach at multiple schools will be assigned the district academic growth average. Student growth will be assessed on matched students by determining the difference between:

- Grades K – 2: the DIBELS or EDL scores are compared from the beginning of the year 2021-22 to the end of the year 2021-22. If these scores cannot be collected, K-2 teachers will receive the school average.
- Grades 3 - 11: the TUSD pre-post scores are compared from the fall results to the mid-spring results in 2021-22. If these scores cannot be collected, grades 3 – 11 teachers will receive the school average.

E. <u>Point Allocation:</u> Teachers will receive a 1 or 1.5 (below average growth or a total of 7 or 10 points), a 2 (average growth or an average of 13 points), or a 2.5 or a 3 (above average growth or an average of 17 or 20 points) that will be added to the Teacher Evaluation points total.
- 'A' teachers with 15 students or more with matched pre-post data will receive their own score.
  - i. Grades K - 2: Scores are for ELA only
  - ii. Grades 3 – 5: Scores are the average of the ELA and Math scores per teacher
  - iii. Grades 6 – 11: Scores are from the subject (ELA or Math) specific to that teacher.

- 'A' teachers in grades 3 – 11 with fewer than 15 students with matched pre-post scores will receive the school subject mean in which they teach.
- 'A' teachers in grades K – 2 with fewer than 15 students with fall and spring DIBELS or EDL scores will receive the DIBELS/EDL school mean.
- 'B' teachers who support math (math interventionist, AP calculus teacher, etc.) will receive the school math mean.
- 'B' teachers who support ELA (literacy specialist, AP English lit, etc.) will receive the school ELA mean.
- 'B' teachers who do not support ELA or math (PE teacher, art teacher, science teacher, etc.) will get the school mean which is a combination of the math and ELA mean.

Teacher Effectiveness Evaluation Model 2021-22

Student Survey

The three Student Surveys are: Grades K-2, Grades 3 – 5, and Grades 6 – 12. Using the Tripod Study from Harvard University as the conceptual foundation, these surveys measure 7 classroom climate constructs including: Care, Challenge, Control, Clarify, Captivate, Confer, and Consolidate. Each survey has a different number of total questions. The K-2 Survey has 10 questions, the 3-5 Survey has 20 questions and the 6-12 Survey has 25 questions. Each of these 3 surveys is scored on a 4-point Likert scale:

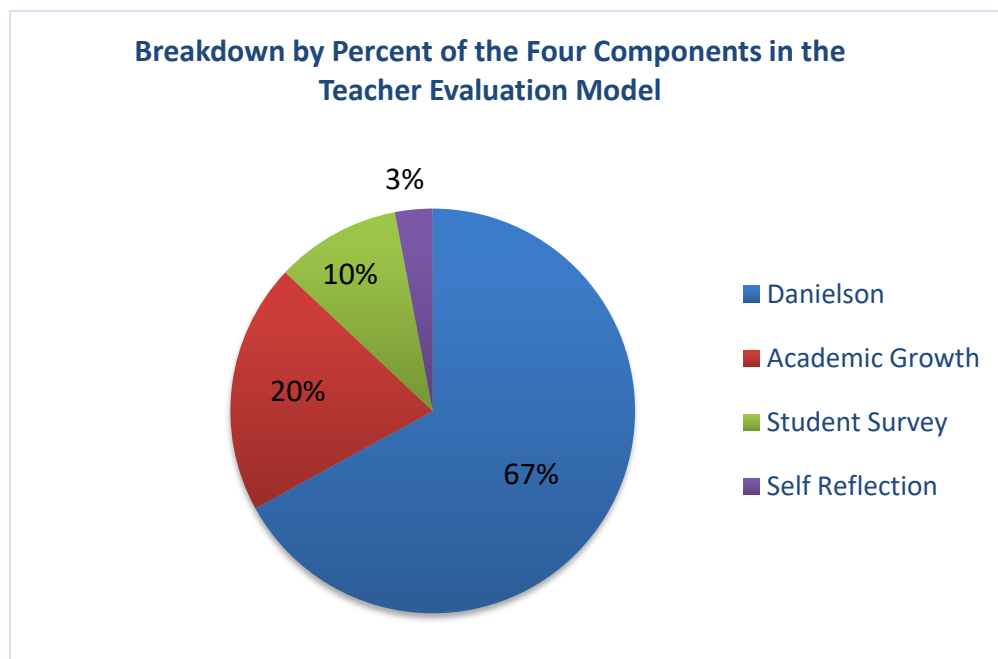|	
1 = Strongly Disagree	3 = Agree
2 = Disagree	4 = Strongly Agree

Responses on the Likert scale are averaged and result in an overall score that ranges from 1 to 4. So, regardless of the grade level and/or number of questions, the score will be the averaged number from the responses.

Teacher Self Reflection

The Teacher Self Reflection is completed by the teacher and is scored either 3 or zero depending on whether it was completed or not.

Converting Raw Scores into Weighted Scores

Breakdown by Percent of the Four Components in the Teacher Evaluation Model

- Danielson — 67%
- Academic Growth — 20%
- Student Survey — 10%
- Self Reflection — 3%

Teacher Effectiveness Evaluation Model 2021-22

Each component of this model carries a different weight as represented in the pie chart above. For example, the results of the Danielson observations are weighted the most heavily because they represent 67% of the total model. The results from the Danielson observations, therefore, will have the greatest impact on a teacher's overall score. Secondly, the academic growth represents 20% of the total model so that it can impact a teacher's overall score, but not necessarily determine the outcome. Finally, the results of the Student Survey (10%) and the Self Reflection Survey (3%) will have a smaller impact on a teacher's overall score.

To get the ratio of the current maximum raw points to desired maximum points, the desired maximum points must be divided by the current raw maximum points. Calculating the ratio using scaling factors will produce properly weighted components.

Because the Desired Maximum Points always add up to 100, it does not matter how many raw maximum points are allocated on the Student Survey or the other components. The scaling factor will always change in response to a change in the maximum raw points of each component so that the weight (Desired Maximum Points) remains constant. See Table 1 for the distribution of points and the scaling factors.

Table 1. Grades K-12 Distribution of Points			
Component	Maximum Raw Points	Scaling Factor*	Desired Max Points
Danielson	88	.761	67
Academic Growth	3	6.67	20
Student Survey	4	2.5	10
Teacher Self Reflection	3	1	3
Total	98		100

* Scaling Factors are derived by dividing the Desired Points by the Maximum Points.

In Tables 2 – 4 below, the raw maximum points are converted into weighted or desired maximum points using a scaling factor. The scaling factor is derived by dividing the Desired Maximum Points (the weighted percent of each component that adds up to 100) by the Current Maximum Raw Points. The scaling factor, therefore, changes the raw points into the weighted points for each component.

Teacher Effectiveness Evaluation Model 2021-22

The following examples show three different Grade 4 teachers with three different sets of raw points. Their points were converted using the Scaling Factor Conversion to give the weighted points.

Teacher A – Grade 4

Table 2. Calculation of Points of a Teacher Scoring about Half of the Possible Points (Developing Teacher Status)			
Component	Raw Points	Scale Conversion	**Weighted Points**
Danielson	44	44 x .761	**33**
Academic Growth	2	2 x 6.67	**13**
Student Survey	2	2 x 2.5	**5**
Teacher Self Reflection	3	3 x 1	**3**
Total	51		**54**

Teacher B – Grade 4

Table 3. Calculation of Points of a Teacher Scoring about Average of the Possible Points (Effective Teacher Status)			
Component	Raw Points	Scale Conversion	**Weighted Points**
Danielson	70	70 x .761	**53**
Academic Growth	2	2 x 6.67	**13**
Student Survey	3.2	3.2 x 2.5	**8**
Teacher Self Reflection	3	1 x 1	**3**
Total	78.2		**77**

Teacher C – Grade 4

Table 4. Calculation of Points of a Teacher Scoring Most Points (High Effective Status)			
Component	Raw Points	Scale Conversion	**Weighted Points**
Danielson	75	75 x .761	**57**
Academic Growth	2	2 x 6.67	**13**
Student Survey	3.2	3.2 x 2.5	**8**
Teacher Self Reflection	3	3 x 1	**3**
Total	83.2		**81**

Teacher Effectiveness Evaluation Model 2021-22

Cut Scores for 2020-21

The cut scores for 2020-21 are:

Ineffective	0 – 46 total points
Developing	47 – 60 total points
Effective	61 – 78 total points
Highly Effective	79 - 100 total points

Based on these cut scores, Teacher A above would be considered "Developing", Teacher B would be considered "Effective", Teacher C would be considered "Highly Effective".

Note: *A methodological improvement was implemented in 2017-18 and will be continued. It provides greater equity of growth scores to 'B' teachers by standardizing the standard deviation and N size of each school to produce an equivalent statistical power across schools. This methodological change results in a reduction of the number of Ineffective (1) and Highly Effective (3) 'B' teachers' growth scores, while not changing the model for 'B' teachers. With a greater number of 'B' teachers receiving the neutral growth score of 2, the final determination of the evaluation relies more heavily on the other components (Danielson Observation by principals, Student Survey of Teachers, Self-Reflection).*