MEMORANDUM July 19, 2017

TO: Gloria Cavazos
Chief Human Resources Officer

Grenita Lathan
Chief Academic Officer

FROM: Carla Stevens
Assistant Superintendent, Research and Accountability

SUBJECT: **TEACHER APPRAISAL AND DEVELOPMENT SYSTEM END OF YEAR REPORT, 2015–2016**

The Houston Independent School District (HISD) Teacher Appraisal and Development System (TADS), launched in 2011–2012, is intended to improve teacher performance and increase student achievement. This report documents teacher appraisal outcomes from the 2015–2016 school year and includes historical data since 2011–2012. Teacher appraisal outcomes are presented by summative ratings across multiple years and by specific teacher and campus level variables. TADS outcomes are then summarized by each appraisal component – Instructional Practice (IP), Professional Expectations (PE), and Student Performance (SP). Finally, this report examines the impact of Student Performance on summative ratings.

Key findings include:

- In 2015–2016, a total 12,255 teachers worked in HISD and were eligible for appraisal under the Teacher Appraisal and Development system. Of these teachers, 11,015 (90%) received an overall TADS summative rating.
- In 2015–2016, 88 percent of teachers were rated as either Effective or Highly Effective in their overall summative appraisal rating, the highest proportion since TADS was introduced in 2011–2012.
- The greatest variation in the proportion of appraisal component ratings between years occurred within Student Performance. With the change in SP measures used in 2015–2016, the proportion of teachers with an SP rating of Level 4 nearly doubled from 35 percent in 2014–2015 to 67 percent in 2015–2016.
- In 2015–2016, the proportion of teachers with a Highly Effective summative rating that included an SP rating (33%) was 11 percentage points higher than for teachers with a Highly Effective summative rating that did not include an SP rating (22%).

Further distribution of this report is at your discretion. Should you have any further questions, please contact me at 713-556-6700.

Attachment

cc: Sam Sarabia
Jeffrey McCanna
Dawn Randle
Sandy Gaw
Abby Taylor

HOUSTON INDEPENDENT SCHOOL DISTRICT

RESEARCH

Educational Program Report

TEACHER APPRAISAL AND DEVELOPMENT SYSTEM END OF YEAR REPORT 2015-2016

HISD | Research and Accountability

ANALYZING DATA, MEASURING PERFORMANCE.

2017 BOARD OF EDUCATION

Wanda Adams
President

Diana Dávila
First Vice President

Jolanda Jones
Second Vice President

Rhonda Skillern-Jones
Secretary

Anne Sung
Assistant Secretary

Anna Eastman
Manuel Rodriguez, Jr.
Michael L. Lunceford
Holly Maria Flynn Vilaseca

Richard A. Carranza
Superintendent of Schools

Carla Stevens
Assistant Superintendent
Department of Research and Accountability

Isabel Hovey
Research Specialist

Lissa Heckelman, Ph.D.
Research Manager

Houston Independent School District
Hattie Mae White Educational Support Center
4400 West 18th StreetHouston, Texas 77092-8501

www.HoustonISD.org

It is the policy of the Houston Independent School District not to discriminate on the basis of age, color, handicap or disability, ancestry, national origin, marital status, race, religion, sex, veteran status, political affiliation, sexual orientation, gender identity and/or gender expression in its educational or employment programs and activities.

Table of Contents

Teacher Appraisal and Development System: End of Year Report, 2015–2016

Executive Summary

Evaluation Description

The Houston Independent School District (HISD) launched the Effective Teachers Initiative (ETI) in 2010 with the goal of increasing teacher effectiveness to improve student achievement. As part of this initiative, HISD implemented the Teacher Appraisal and Development System (TADS) in the 2011–2012 school year to provide teachers, principals, and district officials the information they need to improve teacher performance in the classroom and, thereby, increase student outcomes.

This report documents teacher appraisal outcomes from the 2015–2016 school year and includes historical data since 2011–2012. Teacher appraisal outcomes are presented by summative ratings across multiple years, and by campus level variables and teacher characteristics. Teacher appraisal outcomes are then described by each TADS component – Instructional Practice (IP), Professional Expectations (PE), and Student Performance (SP). Finally, this report examines the impact of Student Performance on summative ratings.

Highlights

- In 2015–2016, 90 percent of teachers (11,015 of 12,255) teachers received an overall TADS summative rating. Of those teachers, 88 percent were rated as either Effective or Highly Effective in their overall summative appraisal rating, the highest proportion since TADS was introduced in 2011–2012. From 2012–2013 to 2015–2016, the mean summative score for all appraised teachers increased from 2.93 to 3.16.

- Met Standard schools had more than double the proportion of Highly Effective teachers (27%) compared to teachers at IR schools (12%). At the same time, teachers in IR schools (23%) were more than twice as likely to be rated as Needs Improvement compared to teachers in Met Standard schools (10%).

- New teachers, with less than one year of experience, were over three times more likely to receive a summative rating of Needs Improvement or Ineffective compared to their more experienced colleagues (34% compared to 10% for all other teachers).

- Teachers were over four times more likely to receive a Level 1 or Level 2 rating for Instructional Practice (IP) than for Professional Expectations (PE) (13% compared to 3%).

- In 2015–2016, by board decision, the district excluded Value-Added and normative assessments from the SP calculations. Consequently, SP ratings for 2015–2016 included only measures of Comparative Growth for TELPAS and Student Progress. With the change in SP measures, the proportion of teachers with an SP rating included in their summative rating fell thirteen percentage points from 2014–2015 to 2015–2016 (43% to 30%). In 2015–2016, teachers were twice as likely to receive a rating of Level 4 for Student Performance (SP) (67%) compared to IP (24% of teachers were rated Level 4) and PE (29% of teachers were rated Level 4).

- The proportion of teachers with a Highly Effective summative rating that included an SP rating (33%) was 11 percentage points higher than the proportion of teachers with a Highly Effective summative rating that did not include an SP rating (22%). At the same time, the proportion of teachers with a Needs Improvement summative rating that included an SP rating (6%) was eight percentage points lower than the proportion of teachers with a Needs Improvement summative rating that did not include an SP rating (14%).

Recommendations

- HISD's teacher appraisal system was designed to provide teachers with consistent, individualized support in order to have the greatest impact on student learning outcomes in the classroom. HISD should consider studying possible correlations between TADS and the facilitation of instructional support to improve teacher effectiveness, possibly by analyzing longitudinal and qualitative data from teachers and appraisers that have participated in TADS across multiple years.

- In efforts to ensure fair support for all teachers, the district should continue to encourage greater TADS participation. One potential area for expanding teacher participation is to include teachers assigned to district charter schools.

- Teachers may not have adequate support to overcome the obstacles and challenges they face at low-performing schools. HISD should continue to focus on growing and supporting teachers located at Improvement Required (IR) schools who typically receive lower ratings than teachers at Met Standard schools. In addition, HISD should continue efforts that attract effective teachers to IR schools. Future research could explore the possible reasons why teachers at IR schools receive lower ratings through analysis of longitudinal and qualitative data for teachers and appraisers that have participated in TADS across multiple years.

- Teachers with less than five years of experience require individualized support focused on improving their effectiveness in the classroom. HISD should continue to support ongoing strategies for these teachers, including efforts to retain effective teachers beyond their first few years of teaching.

- The types of measurements used to calculate the Student Performance component of the TADS system have varied across years. Consequently, two areas of TADS have been impacted by these changes. First, the proportion of teachers that have SP available for their summative rating substantially decreased in 2015–2016. Second, the number of statistically rigorous, objective measures used to calculate SP has decreased, resulting in Student Progress, a subjective, less-rigorous measure, becoming the predominate measure for SP. TADS leadership should continue ongoing efforts to ensure the accuracy and consistency of all student growth measures used to calculate the SP rating (i.e., Value-Added growth, Comparative Growth, and Student Progress).

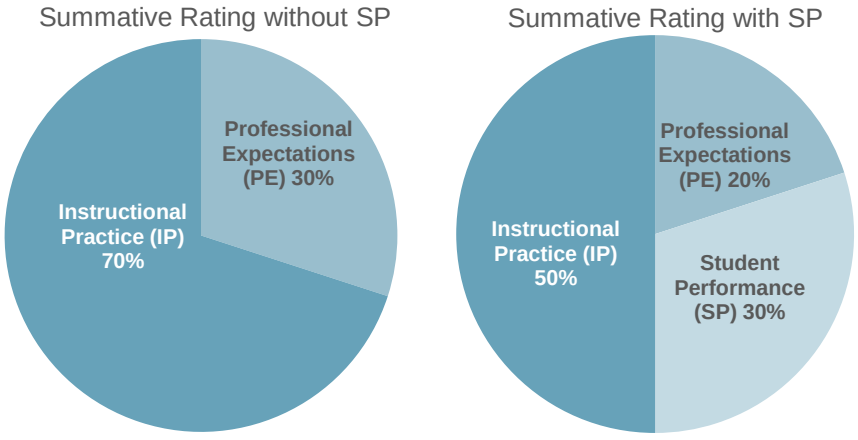
- Findings from this report suggest that the Student Progress measure may increase a teacher's SP score and, consequently, increase overall summative appraisal ratings for teachers that use Student Progress as a measure. Ensuring the accuracy, validity, and stability of the Student Progress measure is an ongoing effort and should remain a focus of TADS leadership.

Introduction

Houston Independent School District's Teacher Appraisal and Development System (TADS), as part of the Effective Teachers Initiative (ETI), is designed to promote effective teaching in every classroom, which ultimately translates into improved student achievement. TADS is intended to give teachers, principals, and district officials the information they need to improve teacher performance in the classroom that ensures the opportunity for every student in the district to learn from an effective teacher (Research and Accountability, 2015). TADS, similar to other well-designed evaluation systems, incorporates multiple, differently-weighted measures of classroom observations and student growth to evaluate teacher effectiveness. A 2016 external audit reported that TADS is statistically comparable to other comprehensive evaluation systems in large districts with available data (Education Analytics, 2016).

In the TADS system, effective teaching is determined by three performance criteria areas, or appraisal components – Instructional Practice (IP), Professional Expectations (PE), and Student Performance (SP). Each appraisal component is based on specific criteria. Further information on the TADS component distribution can be found in **Appendix A** (p.33). Teachers are rated on a scale of one to four for each of the appraisal components. The weighted sum of those appraisal components is then used to calculate a teacher's TADS summative appraisal rating. The components used to calculate a teacher's summative rating vary depending on the measures available to a teacher. Teachers must have at least two measures of student growth or achievement to have SP count in their summative ratings. If a teacher has only one SP measure or no SP measure, the overall TADS summative rating is calculated using 70 percent IP and 30 percent PE ratings. Teachers that receive all three appraisal components (i.e., IP, PE, and SP) receive a summative rating based on 50 percent IP, 20 percent PE, and 30 percent SP (**Figure 1**). A detailed guide of the summative rating components can be found in **Appendix B** (p. 34).

Figure 1. TADS Summative Rating Components Weights, 2015–2016



Source: Student Performance (SP) Guidebook, 2015–2016

For the summative appraisal rating, weighted by available appraisal components, each teacher receives a rating of Ineffective, Needs Improvement, Effective, or Highly Effective. These ratings are scored as: 1.00 to 1.49 – Ineffective, 1.50 to 2.49 – Needs Improvement, 2.50 to 3.49 – Effective, and 3.50 to 4.00 – Highly Effective.

Over the course of the school year, the TADS system pairs each participating teacher with one appraiser. The role of the appraiser is to coach the teacher towards effective teaching practices. Appraisers should

observe teachers to provide useful feedback intended to improve teaching practices and support the teacher in curriculum planning and professional development (Leadership Development, 2013). Appraisers assign scores to the teachers for whom they are responsible using standardized rubrics and support teachers in determining outcome measures. Appraisers use the IP rubric to assess a teacher's skills and ability to promote learning through classroom observations and walkthroughs. Appraisers use the PE rubric to assess a teacher's efforts to meet objective, measurable standards of professionalism. Along with IP and PE, appraisers support teachers through the Student Performance (SP) process, including approving a teacher's student outcome measures. Appendix A (p. 33) presents the various types of weights used to calculate a teacher's SP rating. Further details for each Student Performance measure can be found in **Appendix C** (pp. 35–36).

The goal of this report is to describe the distribution of TADS teacher summative ratings and the performance area appraisal components used to construct a teacher's overall summative appraisal rating. Summative appraisal ratings are described over the five-year duration of HISD's appraisal system for teachers and by comparisons between the 2014–2015 and 2015–2016 school years. This report examines how these ratings were distributed across key variables by campus level and teacher characteristics. Teacher appraisal outcomes are then described by each TADS component – Instructional Practice (IP), Professional Expectations (PE), and Student Performance (SP). Finally, this report examines the potential impact of Student Performance on summative ratings.

Methods

Teacher appraisal data were collected from 2011–2012 to 2015–2016. Ratings for Instructional Practice (IP), Professional Expectations (PE), Student Performance (SP), and overall summative ratings were collected through the TADS Feedback and Development (F&D) Tool and TADS Student Performance (SP) Tool. A teacher was eligible for appraisal if s/he was actively employed from the beginning of the school year until the end of April of each academic year. Teachers not included in the TADS system may have been excluded for a variety of reasons. For example, teachers may not have been rated due to late hiring, job title changes, incorrect job titles in HISD Human Resources Information Systems (HRIS), or split roles that required teachers to teach students less than 50 percent of the instructional day. Moreover, some teachers in leadership roles were appraised in ePerformance, the School Leader Appraisal Tool, rather than in TADS. Finally, teachers employed in HISD charter schools were not appraised in TADS.

For this report, HISD Human Resources (HR) provided district-wide employee rosters, which included multiple identifiers for teacher level data. The ASPIRE (Accelerating Student Progress: Increasing Results and Expectations) team, housed within the HISD Department of Research and Accountability, provided additional identifiers for teacher level data. In each case, only teachers who received a TADS summative rating were included in the analyses. Additional data on school accountability were obtained through the Texas Education Agency (TEA) website. The specific methodology on developing the specific variables used in this report can be found in **Appendix D** (pp. 37–38).

Data Limitations

Changes to the measurements included to calculate the Student Performance TADS component pose a challenge to data analysis and comparisons of TADS ratings over time. **For 2015–2016, these changes refer to exclusion of norm-referenced assessments (Iowa/Logramos) from the SP measure of Comparative Growth and the exclusion of Value-Added Growth as an SP measure.** These two changes had a substantial impact on the comparability with results from prior years.

Data limitations specific to this report include smaller samples of appraised teachers when described by campus and teacher characteristics. Where indicated, the reader will find footnotes explaining data limitations.

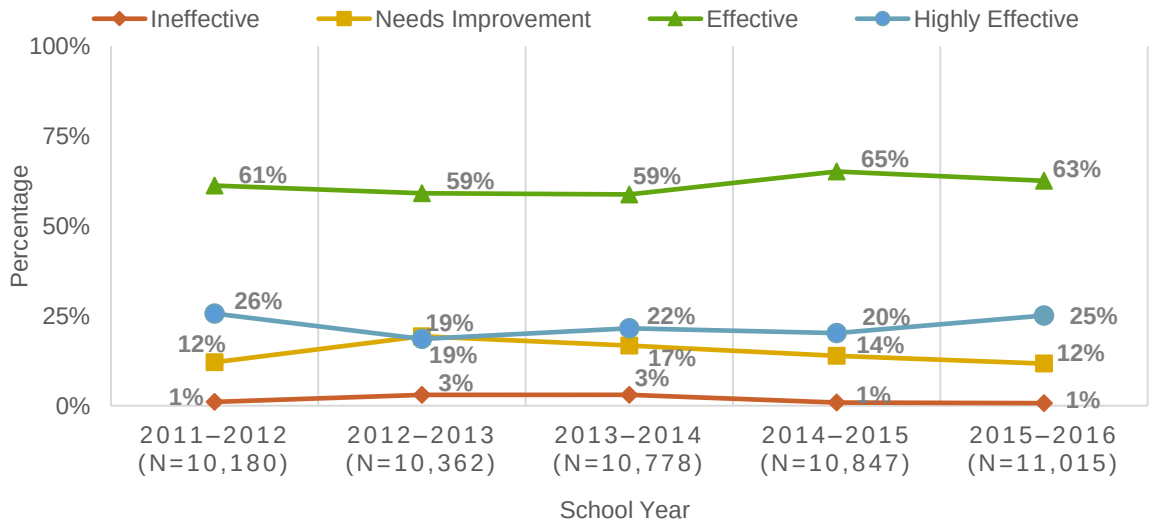
Results

Summative Ratings

What was the distribution of summative ratings for teachers districtwide in 2015–2016 compared to previous years?

- **Figure 2** compares the summative rating distributions from the inception of the TADS system in 2011–2012 through 2015–2016.[1] More details on the summative rating distribution from 2011–2012 to 2015–2016 can be found in **Appendix E** (**Table E-1**, pp. 39–40). The TADS summative rating distribution varied across years. In 2012–2013, the second year of implementation, teachers received the lowest proportion of Effective and Highly Effective summative ratings, when it fell nine percentage points (87% in 2011–2012 to 78% in 2012–2013). Over the following three years, from 2013–2014 to 2015–2016, the proportion of teachers with Effective and Highly Effective ratings increased each year to its highest percentage in 2015–2016 (88%).

Figure 2. Summative Rating Distribution 2011–2012 through 2015–2016



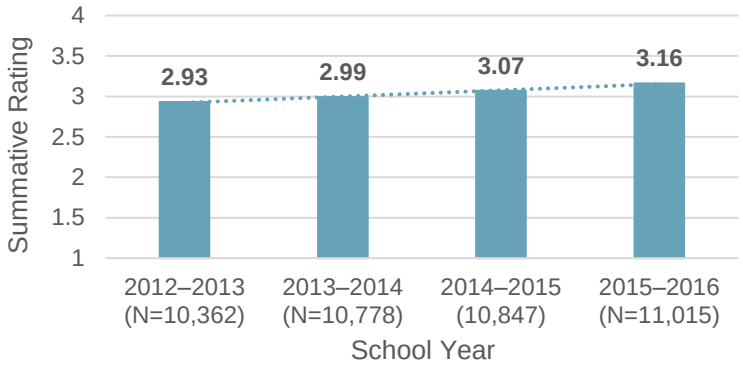
Source: Teacher Appraisal and Development F&D Tool, 2011–2012, 2012–2013, 2013–2014, 2014–2015, 2015–2016
Note: Percentages may not total 100 due to rounding.

- The proportions of teachers with summative ratings of Needs Improvement and Highly Effective have the greatest variation in the distribution from year to year. In 2011–2012, 12 percent of teachers received a Needs Improvement rating. That proportion increased by seven percentage points in 2012–2013 (19%) and then steadily decreased over time to 12 percent again in 2015–2016. In 2011–2012, 26 percent of teachers received a Highly Effective rating. That proportion decreased by seven

[1] In the years preceding 2011–2012, HISD used the state's Professional Development and Appraisal System (PDAS) to appraise its teachers.

percentage points in 2012–2013 (19%) and then increased over time to 25 percent in 2015–2016 (Figure 2).

Figure 3. Mean Summative Score, 2012–2013 through 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2012–2013, 2013–2014, 2014–2015, 2015–2016

- **Figure 3** shows the mean summative scores from 2012–2013 to 2015–2016. The mean summative score for teachers has steadily increased by 0.23 points from 2.93 in 2012–2013 to 3.16 in 2015–2016.

What was the distribution of summative ratings and appraisal components for teachers districtwide in 2015–2016 compared to 2014–2015?

- In 2015–2016, a total 12,255 teachers worked in HISD and were eligible for appraisal under TADS. Of these teachers, 11,015 (90%) received an overall TADS summative rating compared to 10,847 of 11,592 (94%) in the previous 2014–2015 school year.

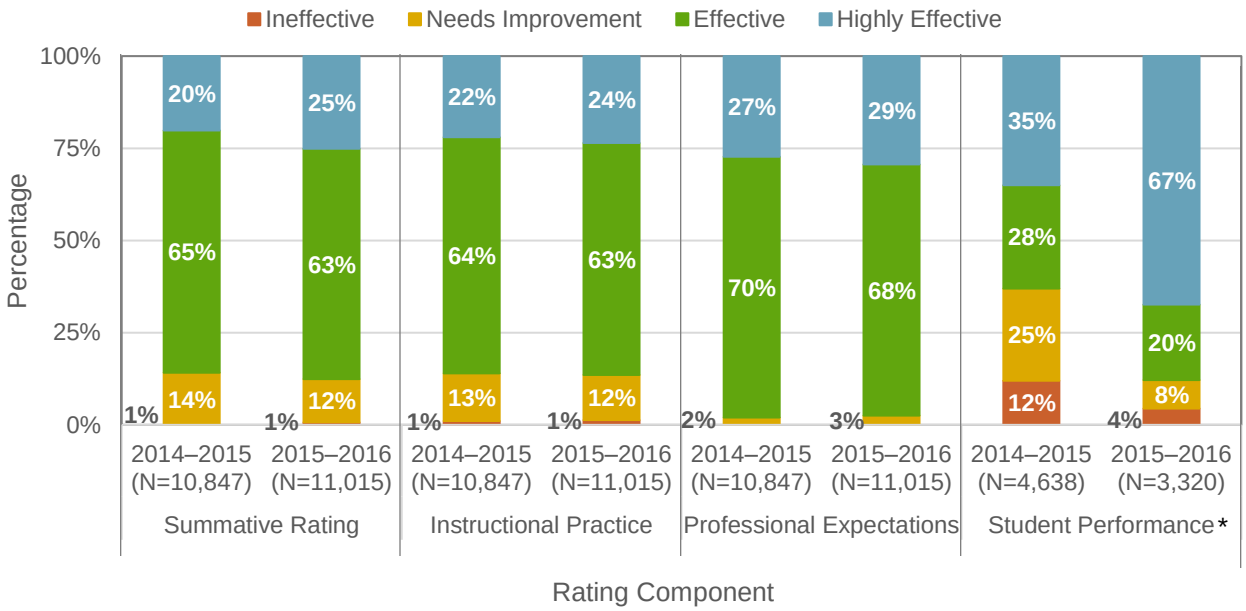
- **Figure 4** (p. 6) shows the distribution of Ineffective, Needs Improvement, Effective, and Highly Effective teachers for 2014–2015 and 2015–2016 by summative rating and appraisal components of Instructional Practice (IP), Professional Expectations (PE), and Student Performance (SP). Of the appraisal components comprising the overall summative rating, both IP and PE remained relatively consistent between years, with marginal increases of two percentage points for teachers receiving Level 4 (i.e., Highly Effective) IP and PE ratings. Conversely, Student Performance (SP) ratings were more variable. More details on the summative rating and appraisal component distribution from 2011–2012 to 2015–2016 can be found in Appendix E (Table E-1, pp. 39–40).

- The greatest variation in the proportion of appraisal component ratings between years occurred within the Student Performance component.[2] Of the 11,015 teachers receiving a summative rating in 2015–2016, 30 percent (n=3,320) received an SP rating, compared to 43 percent (n=4,638) of 10,847 teachers in 2014–2015. From 2014–2015 to 2015–2016, the proportion of teachers with an SP rating of Level 4 (i.e., Highly Effective) nearly doubled from 35 percent in 2014–2015 to 67 percent. The proportion of teachers receiving an SP Level 3 (i.e., Effective) rating fell eight percentage points (28%

[2] The calculation of the Student Performance measures was not consistent between 2014–2015 to 2015–2016. The SP ratings for 2014–2015 included measures of Valued-Added, Comparative Growth for Iowa/Logramos or TELPAS, and Student Progress. The SP ratings for 2015–2016 included measures of Comparative Growth for TELPAS and Student Progress.

to 20%). Finally, the proportion of teachers with an SP Level 1 (i.e., Ineffective) or Level 2 (i.e., Needs Improvement) rating decreased 25 percentage points (37% to 12%) (Figure 4).

Figure 4. TADS Distribution by Summative Rating and by Appraisal Component, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; Teacher Appraisal and Development SP Tool, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding. *See Data Limitations on p. 4.

- **Table 1** (p. 8) shows performance level changes of summative ratings for teachers who received a summative rating in both 2014–2015 and 2015–2016. Of the 11,015 teachers that received a summative rating for 2015–2016, 78 percent of teachers (n=8,634) also received a summative rating for the previous year, 2014–2015.

- The mean summative score for the 8,634 teachers with summative ratings both years increased by 0.12 points from 2014–2015 to 2015–2016. A paired t-test[3] was conducted to compare the mean summative scores received by teachers that received a summative rating in both 2014–2015 and 2015–2016. The mean summative score for 2015–2016 was significantly higher (M = 3.23, SD = 0.51) compared to 2014–2015 (M = 3.11, SD = 0.50), t (8,633) = 23.37, p<0.01.

- A total of 5,787 teachers rated as Effective in 2014–2015 received a summative rating again in 2015–2016. Of those teachers, nearly 18 percent of them (n=1,031) increased their summative rating to Highly Effective in 2015–2016 (Table 1, p. 8).

[3] A paired t-test is a statistical procedure to determine whether the difference in means between two groups is significant or due to random chance.

Table 1. Summative Ratings Changes from 2014–2015 to 2015–2016 for Teachers Who Received a Summative Rating in Both Years (N=8,634)					
	2014–2015 Summative Ratings				
2015–2016 Summative Ratings	Ineffective	Needs Improvement	Effective	Highly Effective	**Total in 2015–2016**
Ineffective	6	20	13	1	**40**
Needs Improvement	8	287	368	12	**675**
Effective	4	640	4,375	399	**5,418**
Highly Effective	0	42	1,031	1,428	**2,501**
Total in 2014–2015	**18**	**989**	**5,787**	**1,840**	**8,634**

Source: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016

Notes: Percentages may not total 100 due to rounding. Red represents 1) teachers receiving an Ineffective rating both years and 2) teachers who fell to Ineffective rating or Needs Improvement rating in 2015–2016 from a higher rating in 2014–2015. Yellow represents teachers 1) remaining as Needs Improvement both years and 2) teachers who fell from Highly Effective in 2014–2015 to Effective in 2015– 2016. Green represents 1) teachers who increased their ratings from 2014–2015 to 2015–2016 and 2) teachers remaining as Effective or Highly Effective.

- Of the 989 teachers rated as Needs Improvement in 2014–2015, nearly 69 percent (n=682) increased their summative rating to Effective or Highly Effective in 2015–2016 (Table 1).

- For comparison, **Table 2** shows performance level changes of summative ratings for teachers who received a summative rating in both 2013–2014 and 2014–2015. Of the 10,847 teachers that received a summative rating for 2014–2015, 77 percent of teachers (n=8,313) also received a summative rating for the previous year, 2013–2014.

Table 2. Summative Ratings Changes from 2013–2014 to 2014–2015 for Teachers Who Received a Summative Rating in Both Years (N=8,313)					
	2013–2014 Summative Ratings				
2014–2015 Summative Ratings	Ineffective	Needs Improvement	Effective	Highly Effective	**Total in 2014–2015**
Ineffective	6	15	10	0	**31**
Needs Improvement	67	328	347	25	**767**
Effective	49	823	3,951	700	**5,523**
Highly Effective	2	62	743	1,185	**1,992**
Total in 2013–2014	**124**	**1,228**	**5,051**	**1,910**	**8,313**

Source: Teacher Appraisal and Development F&D Tool, 2013–2014 and 2014–2015

Notes: Percentages may not total 100 due to rounding. Red represents 1) teachers receiving an Ineffective rating both years and 2) teachers who fell to Ineffective rating or Needs Improvement rating in 2014–2015 from a higher rating in 2013–2014. Yellow represents teachers 1) remaining as Needs Improvement both years and 2) teachers who fell from Highly Effective in 2013–2014 to Effective in 2014–2015. Green represents 1) teachers who increased their ratings from 2013–2014 to 2014–2015 and 2) teachers remaining as Effective or Highly Effective.

- Of 5,051 teachers rated as Effective in 2013–2014 who received a summative rating the following year, 15 percent of them (n=743) increased their summative rating to Highly Effective in 2014–2015 (Table 2). This was three percentage points lower than the 18 percent of teachers who similarly increased their summative rating from 2014–2015 to 2015–2016 (Table 1, p. 8).

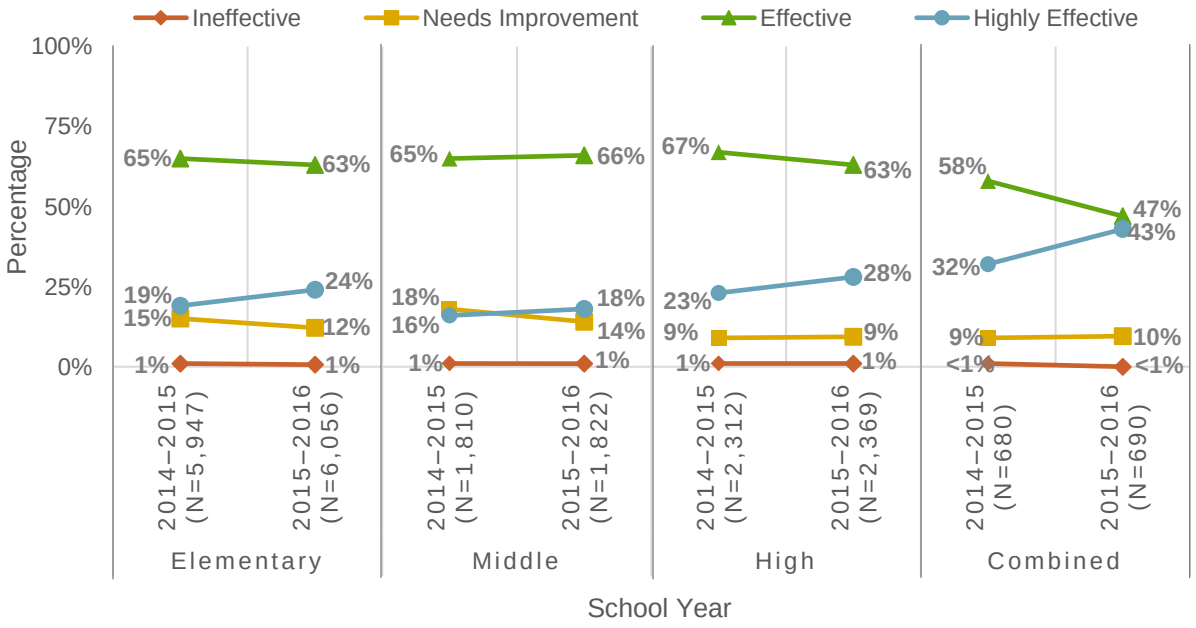
- In 2015–2016, the proportion of teachers that fell from a rating of Highly Effective to Effective decreased by 15 percentage points compared to the previous years. From 2013–2014 to 2014–2015, 37 percent of teachers who received a Highly Effective summative rating in 2013–2014 (n=1,910) fell to an Effective summative rating in 2014–2015 (n=700) (Table 2, p. 8). From 2014–2015 to 2015–2016, 22 percent of teachers who received a Highly Effective summative rating in 2014–2015 (n=1,840) fell to an Effective summative rating in 2015–2016 (n=399) (Table 1).

What was the distribution of summative ratings for teachers by groups in 2015–2016 compared to 2014–2015?

School Level

- **Figure 5** shows summative rating distributions by school level – Elementary, Middle, High, and Combined School levels. More details on the summative rating distribution by school level can be found in Appendix E (Table E-1, pp. 39–40). Consistent with overall summative ratings, the proportion of teachers at all school levels rated Highly Effective increased from 2014–2015 to 2015–2016. The largest increase in the proportion of teachers with a Highly Effective rating occurred at the Combined School level, with an 11 percentage point increase from 2014–2015 to 2015–2016 (32% to 43%).

Figure 5. Summative Rating Distribution by School Level, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Notes: Percentages may not total 100 due to rounding. Teachers located at Community Services, HCC Life Skills, EL DAEP, and Beechnut Academy (n=89) are not included. Teachers without HR identifying information (n=35) are not included.

- Consistent with overall summative ratings, the proportion of teachers at most school levels rated Effective decreased from 2014–2015 to 2015–2016, with the exception of Middle Schools, where a one percentage point increase was observed. The largest decrease in the proportion of teachers with an Effective rating was at the Combined School level, with an eleven percentage point decrease from 2014–2015 to 2015–2016 (58% to 47%) (Figure 5, p. 9).

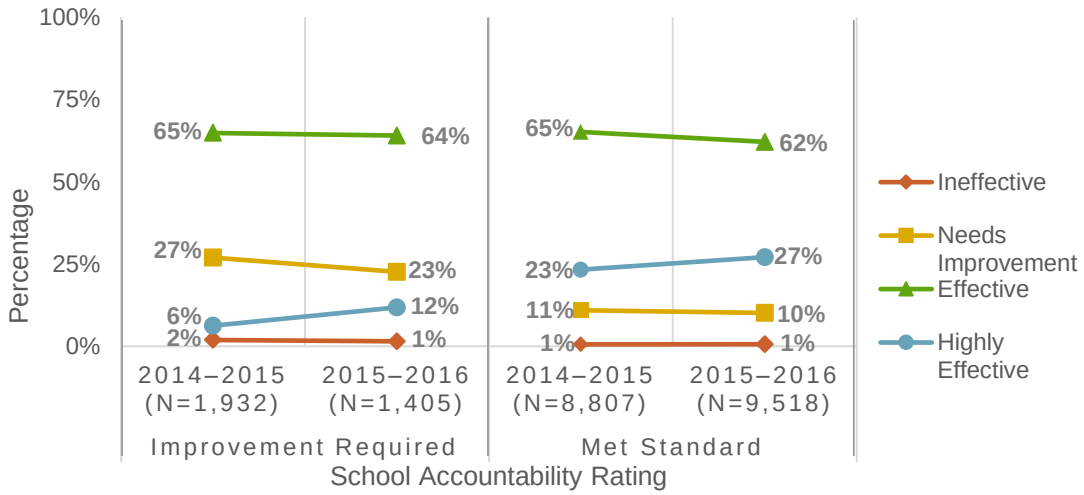
- In 2015–2016, the distribution of summative ratings for teachers at Combined Level schools was inconsistent with the distribution for teachers at other school levels. In 2015–2016, the proportion of teachers rated Effective (47%) was at least sixteen percentage points lower than the proportions of teachers rated as Effective at other school levels. Moreover, in 2015–2016, the proportion of teachers rated Highly Effective (43%) was at least fifteen percentage points higher than the proportions of teachers rated Highly Effective at other school levels (Figure 5).

- When separated by school level, there was high variation in the mean summative score across school levels. Teachers at the Combined School level received the highest mean summative score, 3.34, and teachers at the Middle School level received the lowest mean summative score, 3.06. The mean summative score for each school level remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative scores by school level can be found in **Appendix F** (**Table F-1**, p. 43).

Accountability Rating

- **Figure 6** illustrates the summative ratings distribution by schools' accountability rating of Improvement Required (IR) or Met Standard for 2014–2015 and 2015–2016. More details on the summative rating distribution by accountability rating can be found in Appendix E (Table E-1, p. 39–40). Consistent with overall summative ratings, both accountability rating groups showed an increase in the proportion of teachers rated Highly Effective from 2014–2015 to 2015–2016.

Figure 6. Summative Rating Distribution by School Accountability Rating, 2014–2015 and 2015– 2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; TEA Accountability, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. Teachers located at Not Rated (NR) schools (n=57) are not included. Teachers without HR identifying information (n=35) are not included.

- The separation of teachers by their 2015–2016 campus accountability rating showed a gap of 15 percentage points between the proportion of Highly Effective Teachers at Met Standard and IR schools. Specifically, Met Standard schools had more than double the proportion of Highly Effective teachers (27%) compared IR schools (12%) (Figure 6, p. 10).

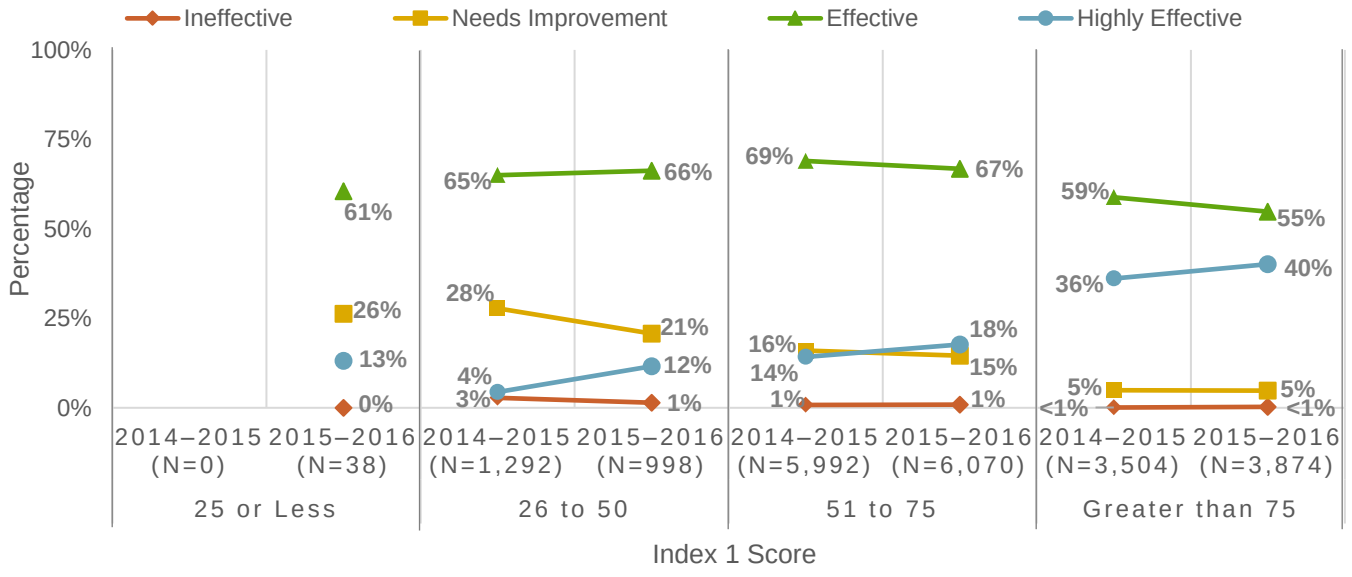
- The proportion of teachers appraised as Needs Improvement in IR schools was higher than that of teachers in Met Standard schools. In 2015–2016, teachers in IR schools (23%) were more than twice as likely to be rated as Needs Improvement compared to teachers in Met Standard schools (10%) (Figure 6).

- Both the mean summative score for Met Standard schools, 3.19, and Improvement Required schools, 2.93, remained within the range required for an Effective summative rating (2.50 to 3.49). The mean summative score for teachers at Met Standard schools was significantly higher compared to the mean summative score teachers at IR schools. More details on the mean summative scores by accountability rating can be found in Appendix F (**Table F-2**, p. 43).

Index 1 Score

- **Figure 7** shows summative ratings by school Index 1 scores for 2014–2015 and 2015–2016. Index 1 scores are an indication of student achievement at the school level, based on student performance on the STAAR test. More details on the summative rating distribution by Index 1 score can be found in Appendix E (Table E-1, p. 39–40). Consistent with overall summative ratings from 2014–2015 to 2015–2016, there was an increase in the proportion of teachers rated as Highly Effective across all groups except those with an Index 1 score of 25 or less.

Figure 7. Summative Rating Distribution by School Index 1 Score Groups, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; TEA Accountability, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

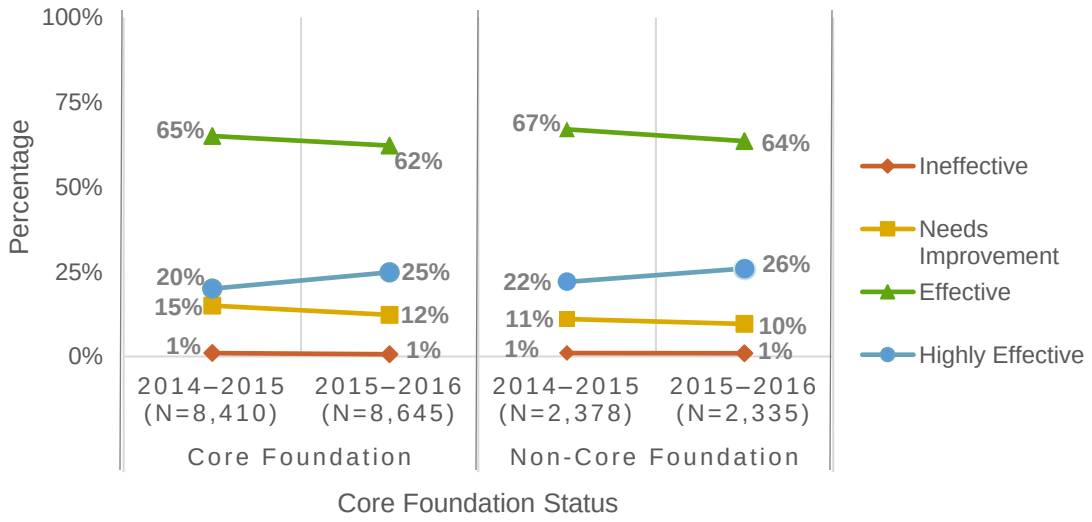
Notes: Percentages may not total 100 due to rounding. In 2014–2015, there were no schools with an Index 1 score of less than 25. In 2015–2016, the schools receiving an Index 1 score of less than 25 (n=38) had an Index 1 Target score of 35. Teachers without HR identifying information (n=35) are not included.

- When separated by 2015–2016 Index 1 scores, there was high variation in the mean summative score across Index 1 score groups. The teachers at schools with Index 1 scores less than 25 received the lowest mean summative score, 2.91, and the teachers at schools with Index 1 scores greater than 75 received the highest mean summative score, 3.36. The mean summative scores by Index 1 score groups remained within the range required for a summative rating of Effective (2.50 to 3.49). More details on the mean summative score by Index 1 score can be found in Appendix F (**Table F-3**, p. 43).

Core Foundation

- **Figure 8** displays summative ratings by Core Foundation teachers and Non-Core Foundation teachers for 2014–2015 and 2015–2016. More details on the summative rating distribution by Core Foundation Status can be found in Appendix E (Table E-1, p. 39–40). In 2015–2016, a slightly higher proportion of Non-Core Foundation teachers were rated Effective and Highly Effective compared to Core Foundation teachers (90% compared to 87%). Accordingly, a slightly lower proportion of Non-Core Foundation teachers were rated Needs Improvement and Ineffective compared to Core Foundation teachers (11% compared to 13%).

Figure 8. Summative Rating Distribution by Core Foundation Teacher Status, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. Teachers without HR identifying information (n=35) are not included.

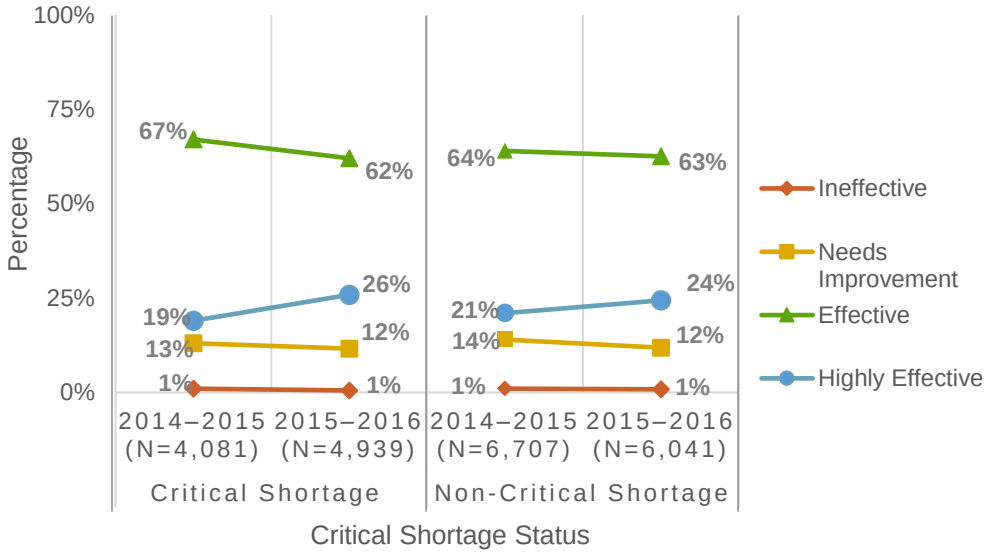
- In 2015–2016, the mean summative score for Non-Core Foundation teachers, 3.19, was higher compared to the score for Core Foundation teachers, 3.15. The mean summative score by Core Foundation status remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative scores by Core Foundation status can be found in Appendix F (**Table F-4**, p. 44).

Critical Shortage

- **Figure 9** (p.13) displays summative ratings by critical shortage and non-critical shortage status for 2014–2015 and 2015–2016. More details on the summative rating distribution by critical shortage status can be found in Appendix E (Table E-1, p.39–40). Critical shortage teachers and non-critical shortage teachers had similar proportions of summative rating performance levels. In 2015–2016, 88 percent of

critical shortage teachers and 87 percent of non-critical shortage teachers were rated Highly Effective or Effective.

Figure 9. Summative Rating Distribution by Critical Shortage Teacher Status, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding. Teachers without HR identifying information (n=35) are not included.

- Consistent with overall summative ratings, the proportion of teachers rated Highly Effective for both critical shortage and non-critical shortage groups increased in 2015–2016 compared to 2014–2015. Moreover, the proportion of teachers rated as Effective or Needs Improvement decreased slightly and the proportion of teachers rated as Ineffective remained stable (Figure 9).

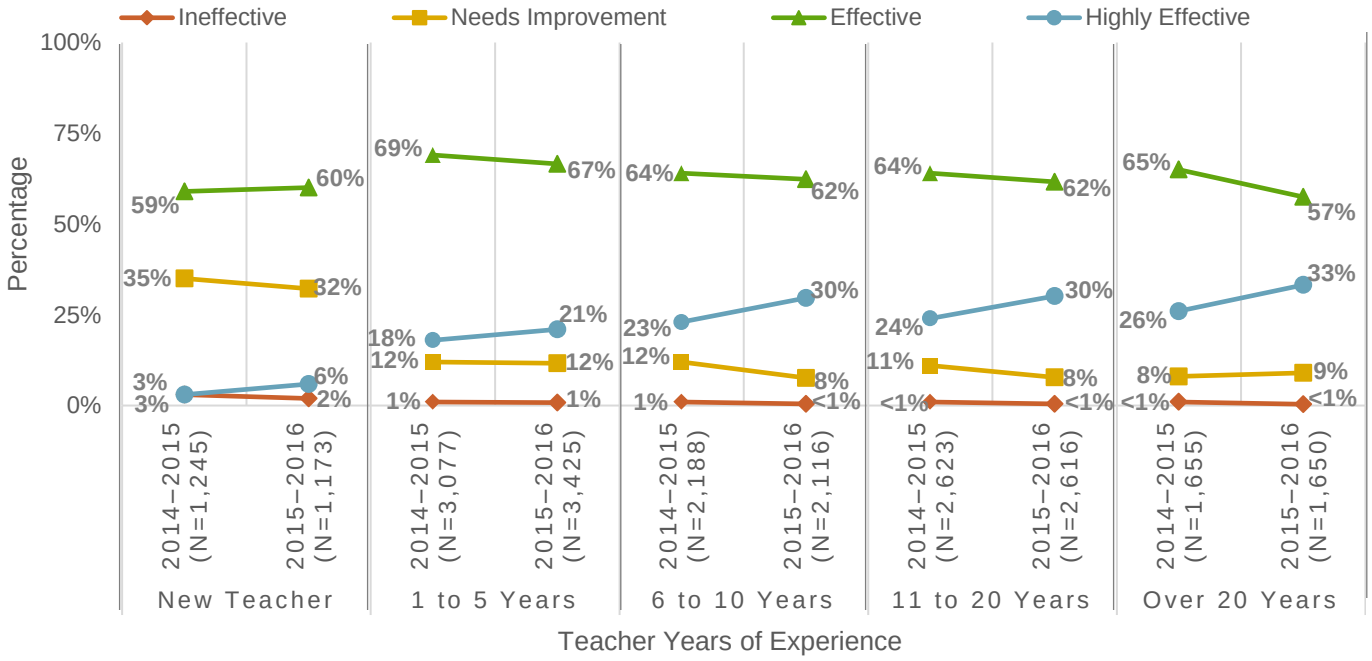
- The mean summative score for critical shortage teachers, 3.17, was higher compared to the score for non-critical shortage teachers, 3.14. The mean summative scores by critical shortage status remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative score by critical shortage status can be found in Appendix F (**Table F-5**, p. 44).

Teachers' Total Years of Experience

- **Figure 10** (p.14) illustrates the summative rating distribution by teachers' total years of experience for 2014–2015 and 2015–2016. More details on the summative rating distribution by teachers' total years of experience can be found in Appendix E (Table E-1, p. 39–40). In 2015–2016, the majority of teachers within each group by total years of experience were rated Effective. Teachers with one to five years of experience had the highest proportion of teachers rated Effective (67%).

- Consistent with overall summative ratings, the proportion of teachers at all experience group levels rated Highly Effective increased from 2014–2015 to 2015–2016. The greatest increase in the proportion of teachers with a Highly Effective rating occurred for teachers with 6 to 10 years and over 20 years of experience, with a seven percentage point increase from 2014–2015 to 2015–2016 (23% to 30% and 26% to 33%, respectively). Teachers with over 20 years of experience had the highest proportion of teachers rated as Highly Effective in 2015–2016 (33%) (Figure 10).

Figure 10. Summative Rating Distribution by Teacher Total Years of Experience, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. Teachers without HR identifying information (n=35) are not included.

- New teachers, with less than one year of experience, overall were more than three times more likely to be rated Ineffective or Needs Improvement compared to their more experienced colleagues (34% compared to 10% for all other teachers[4]) (Figure 10). The data showing new teachers as having a higher proportion rated as Ineffective or Needs Improvement compared to more experienced teachers has held true for the last four school years as well. For a comparison of previous years, see Appendix E (Table E-1, p. 39–40).

- Teachers with over 20 years of experience saw the greatest change of summative ratings distributions from 2014–2015 to 2015–2016. The proportion of teachers rated as Effective decreased by eight percentage points (65% to 57%) and the proportion of teachers rated as Highly Effective increased by seven percentage points (26% to 33%) (Figure 10).

- There was high variation in the mean summative scores across groups of teachers' total years of experience. New teachers, with less than one year of experience, received the lowest mean summative score, 2.79. Teachers with 1 to 5 years of experience received a mean summative score of 3.12. Teachers with 6 to 10 years, 11 to 20 years, and more than 20 years of experience received similar mean summative scores (3.23, 3.24, and 3.26, respectively). The mean summative scores by total years of experience remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative score by teachers' total years of experience can be found in Appendix F (**Table F-6**, p. 44).

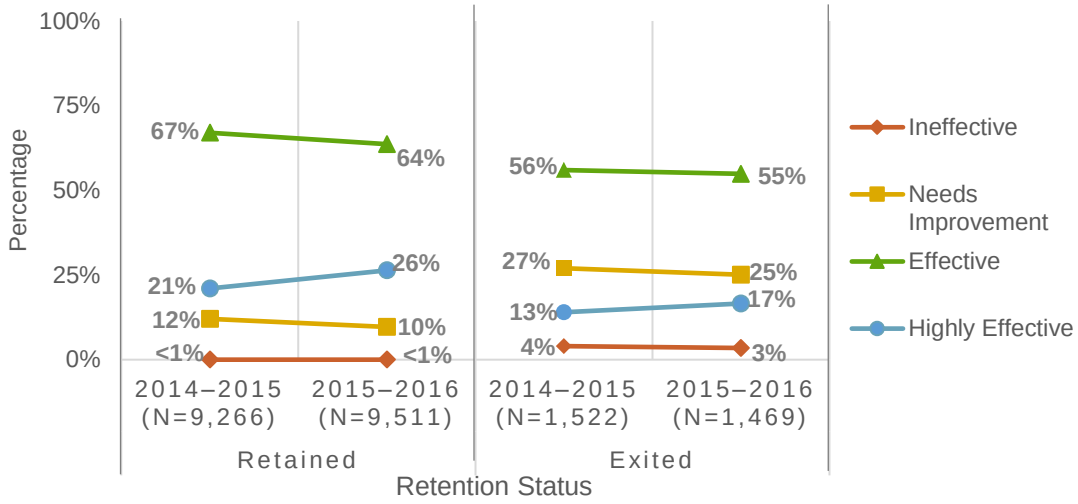
[4] Calculated as the percentage of teachers who were not new who received an Ineffective or Needs Improvement rating (966 out of 9,807 teachers).

Retention

- **Figure 11** displays teacher retention by summative ratings for 2014–2015 and 2015–2016. More details on the summative rating distribution by teacher retention can be found in Appendix E (Table E-1, p. 39–40). Regarding total retention for 2015–2016 to 2016–2017, 87 percent of teachers that received a summative rating during the 2015–2016 school year (9,511 of 10,980) remained in HISD.

- The majority of the 2015–2016 teachers retained for the 2016–2017 school year (90%) received an Effective or Highly Effective summative rating in 2015–2016. (Figure 11).

Figure 11. Summative Rating Distribution by Teacher Retention, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. In 2014–2015, there were 59 teachers without HR identifying information. In 2015–2016, there were 35 teachers without HR identifying information.

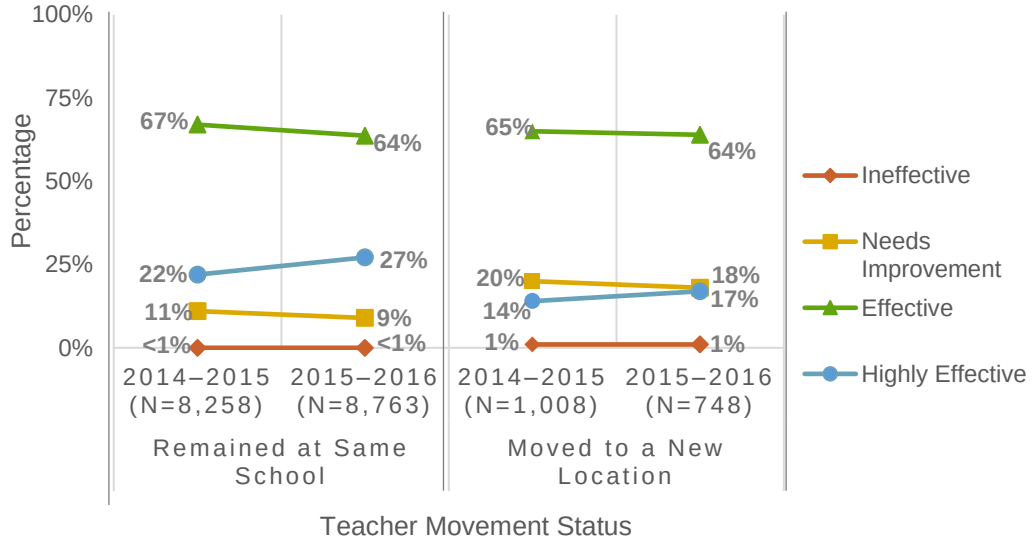
- In 2015–2016, over one quarter (28%) of teachers who exited HISD had a Needs Improvement or Ineffective summative rating, a decrease of three percentage points from the previous school year. While 28 percent of exited teachers left with an Ineffective or Needs Improvement rating, the majority of teachers that left (72%) had an Effective or Highly Effective rating (Figure 11).

- The mean summative score for teachers that chose to remain with HISD, 3.19, was higher compared to teachers that exited the district, 2.90. The mean summative scores by retention status remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative rating by retention can be found in Appendix F (**Table F-7**, p. 45).

Teacher Movement

- **Figure 12** (p.16) describes teacher movement by summative ratings. More details on the summative rating distribution by teacher movement can be found in Appendix E (Table E-1, p. 38–39). Regarding teacher movement, 92 percent of teachers (n=8,763) who received a summative rating during the 2015–2016 school year and remained in HISD (n=9,511) did not change locations within HISD for the following 2016–2017 school year.

Figure 12. Summative Rating Distribution by Teacher Movement, 2014–2015 to 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding.

- Overall, the proportion of teachers who moved to a new location received lower summative ratings than teachers that remained at their 2014–2015 school. In 2015–2016, teachers who moved to a new location had a higher proportion of Ineffective and Needs Improvement ratings than teachers who remained at the same school (19% compared to 10%). The proportion of teachers rated Highly Effective who had moved to a new location (17%) was 10 percentage points lower than for teachers that remained at their same school (27%). However, the proportion of teachers rated Effective was equal for both groups in 2015–2016 (Figure 12).

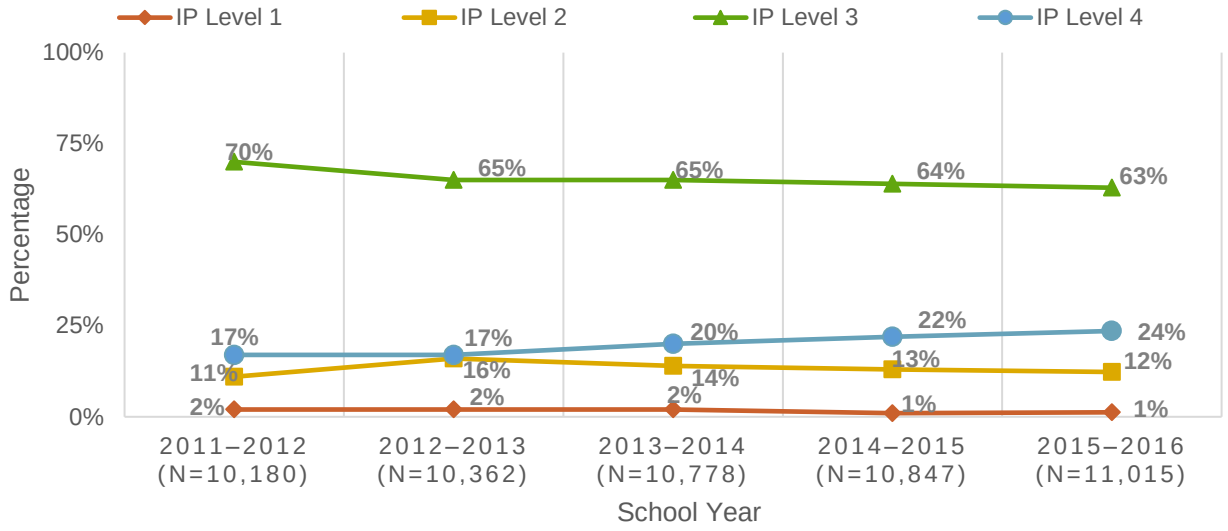
- The mean 2015–2016 summative score for teachers that remained at the same school, 3.21, was higher compared to the mean score of teachers that moved to a new location, 3.03. The mean summative scores by teacher movement status remained within the range required for an Effective summative rating (2.50 to 3.49). More details on the mean summative score by teacher movement can be found in Appendix F (**Table F-8**, p. 45)

Instructional Practice Ratings

What was the distribution of Instructional Practice (IP) ratings in 2015–2016 compared to previous years?

- **Figure 13** (p.17) shows the IP ratings distribution from 2011–2012 through 2015–2016. Instructional Practice ratings have remained relatively stable. More details of the IP rating distribution from 2011–2012 to 2015–2016 can be found in Appendix E (**Table E-2**, p. 41). Across the five school years, the majority of teachers received a Level 3 IP rating. From 2011–2012 to 2015–2016, the proportion of teachers earning a Level 3 IP rating decreased seven percentage points. In 2015–2016, 63 percent of teachers were rated Level 3, a one percentage point decrease from 2014–2015.

Figure 13. Instructional Practice (IP) Rating Distribution 2011–2012 through 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2011–2012, 2012–2013, 2-13–2014, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding.

- The proportion of teachers earning a Level 4 IP rating increased seven percentage points from 2011–2012 to 2015–2016. From 2014–2015 to 2015–2016, there was an increase of two percentage points in the proportion of teachers rated Level 4, from 22 percent to 24 percent. This proportion has increased steadily since 2011–2012 and 2012–2013, when it was 17 percent (Figure 13).

- **Table 3** shows the IP rating changes for teachers who received a IP rating in both 2014–2015 and 2015–2016. Of the 11,015 teachers that received an IP rating for 2015–2016, 78 percent (n=8,634) of teachers received an IP rating for both 2014–2015 and 2015–2016.

Table 3. Instructional Practice (IP) Changes for Teachers with Consecutive IP Ratings, 2014–2015 through 2015-2016 (N=8,634)

	2014–2015 Summative Ratings				
2015–2016 Summative Ratings	IP Level 1	IP Level 2	IP Level 3	IP Level 4	**Total in 2015–2016**
IP Level 1	12	31	25	1	**69**
IP Level 2	17	314	373	10	**714**
IP Level 3	13	578	4,465	441	**5,497**
IP Level 4	1	20	845	1,488	**2,354**
Total in 2014–2015	**43**	**943**	**5,708**	**1,940**	**8,634**

Source: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016
Notes: Percentages may not total 100 due to rounding. Red represents 1) teachers receiving an IP Level 1 rating both years and 2) teachers who fell to IP Level 1 or IP Level 2 in 2015–2016 from a higher rating in 2014–2015. Yellow represents teachers 1) remaining as IP Level 2 both years and 2) teachers who fell from IP Level 4 in 2014–2015 to IP Level 3 in 2015–2016. Green represents 1) teachers who increased their ratings from 2014–2015 to 2015–2016 and 2) teachers remaining as IP Level 3 or IP Level 4.

- A total of 5,708 teachers rated as IP Level 3 in 2014–2015 received an IP rating again in 2015–2016. Of those teachers, nearly 15 percent of them (n=845) increased their IP rating to Level 4 and 78 percent (n=4,465) maintained a Level 3 rating in 2015–2016 (Table 3, p. 17).

- Of the 943 teachers rated as Needs Improvement in 2014–2015, 63 percent (n=598) increased their IP rating to Level 3 or Level 4 in 2015–2016 (Table 3).

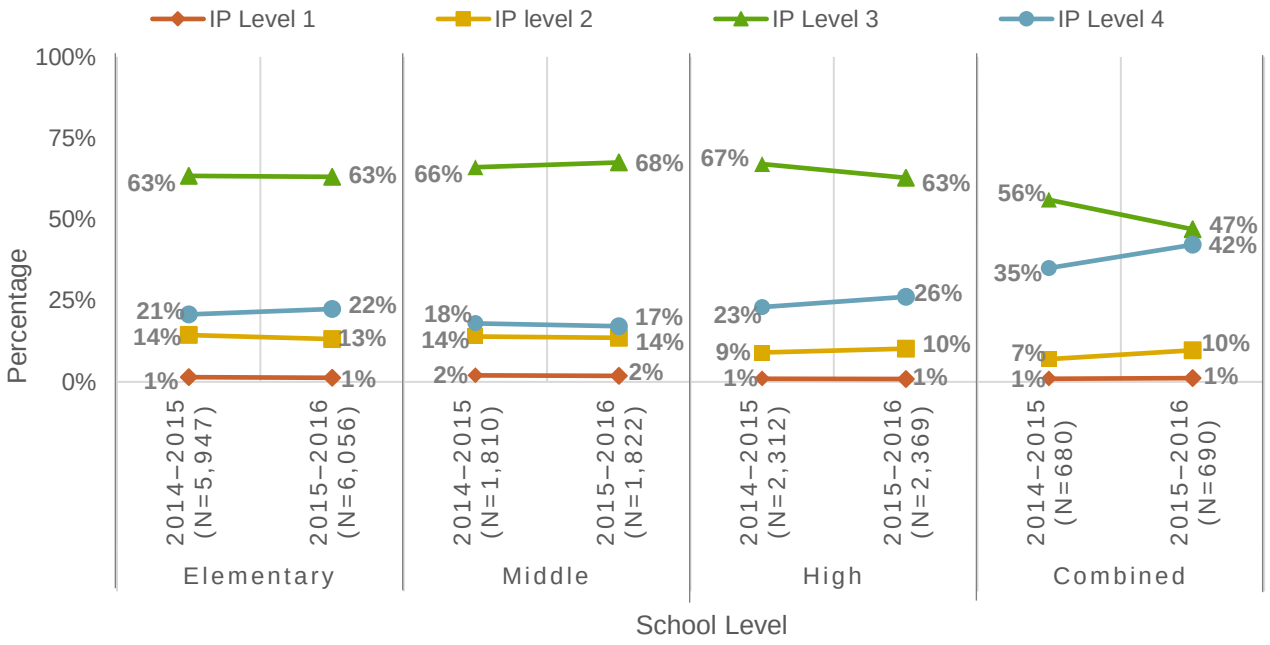
- A paired t-test[5] was conducted to compare the mean IP score received by teachers that received an IP rating in both 2014–2015 and 2015–2016. The mean IP score for the 8,634 teachers appraised both years increased by 0.06 points from 2014–2015 to 2015–2016. The mean IP score for 2015–2016 was significantly higher (M = 3.17, SD = 0.61) compared to 2014–2015 (M = 3.11, SD = 0.59), t (8,633) = 11.74, p<0.01.

What was the distribution of Instructional Practice (IP) ratings in 2015–2016 compared to 2014–2015 and by groups?

School Level

- **Figure 14** shows IP ratings by school level for 2014–2015 and 2015–2016. More details on the IP rating distribution by school level can be found in Appendix E (Table E-2, p. 41). The proportion of teachers at Elementary, High School, and Combined school levels with an IP Level 4 rating increased from 2014–2015 to 2015–2016.

Figure 14. Instructional Practice (IP) Rating Distribution by School Level, 2014–2015 and 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not equal 100 due to rounding. Teachers located at Community Services, HCC Life Skills, EL DAEP, and Beechnut Academy (n=89) are not included. Teachers without HR identifying information (n=35)

[5] A paired t-test is a statistical procedure to determine whether the difference in means between two groups is significant or due to random chance.

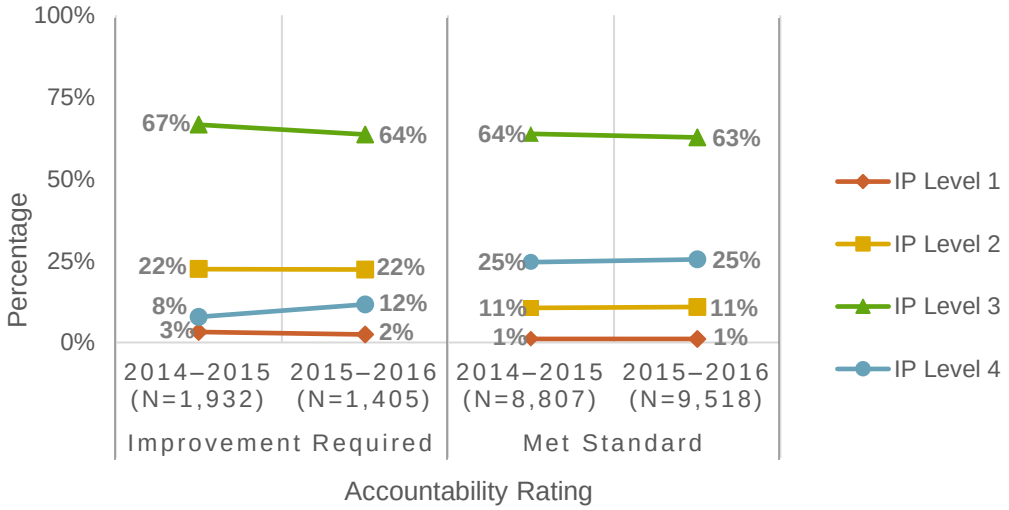
- In 2015–2016, the distribution of IP ratings for teachers at Combined Level schools was inconsistent with the distribution of IP ratings for teachers at other school levels. In 2015–2016, the proportion of teachers rated IP Level 3 (47%) was at least sixteen percentage points lower than the proportions of teachers rated IP Level 3 at other school levels. At the same time, for 2015–2016, the proportion of teachers rated Level 4 (42%) was at least sixteen percentage points higher than the proportions of teachers rated IP Level 4 at other school levels (Figure 14, p. 18).

Accountability Rating

- **Figure 15** displays teachers' IP ratings by school accountability rating for 2014–2015 and 2015–2016. More details on the IP rating distribution by accountability rating can be found in Appendix E (Table E-2, p. 41). Similar to the summative rating distribution by accountability rating (Figure 6, p. 10), the separation of teachers by their campus accountability rating showed a gap between Met Standard and IR schools. In 2015–2016, Met Standard schools had more than double the proportion of IP Level 4 teachers (25%) compared to teachers at IR schools (12%).

- In 2015–2016, the proportion of teachers appraised as Level 2 in IR schools was higher than that of teachers in Met Standard schools. Teachers in IR schools (22%) were twice as likely to receive an IP Level 2 rating compared to teachers in Met Standard schools (11%) (Figure 15).

Figure 15. Instructional Practice (IP) Rating Distribution by School Accountability Rating, 2014–2015 and 2015–2016



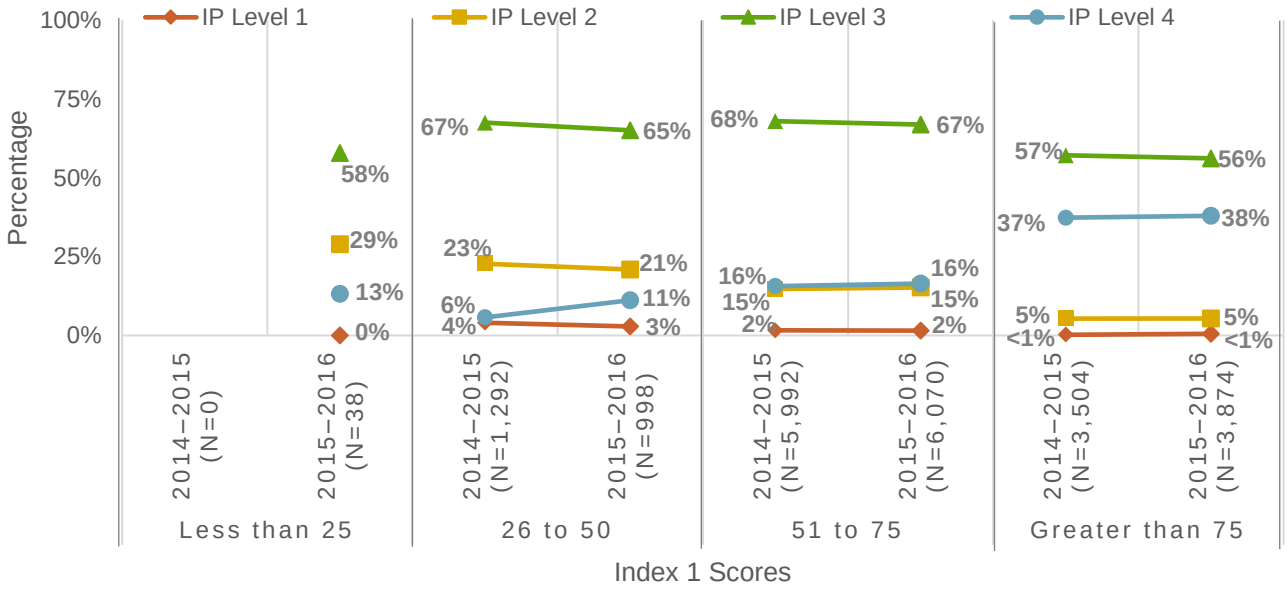
Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016; TEA Accountability Ratings, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding. Teachers located at Not Rated (NR) schools (n=57) are not included. Teachers without HR identifying information (n=35) are not included.

- From 2014–2015 to 2015–2016, the proportion of teachers receiving an IP Level 3 decreased slightly in both groups. At IR schools, the proportion of teachers with a Level 3 rating decreased three percentage points (67% to 64%), and at Met Standard schools, the proportion of teachers with a Level 3 rating decreased one percentage point (64% to 63%) (Figure 15).

Index 1 Score

- **Figure 16** illustrates teachers' IP ratings across categories of schools' Index 1 scores for 2014–2015 and 2015–2016. More details on the IP rating distribution by Index 1 score can be found in Appendix E (Table E-2, p. 41). Similar to the distribution of summative ratings by Index 1 score group (Figure 7, p. 11), the greatest change in the IP rating distribution by Index 1 score from 2014–2015 to 2015–2016 was for teachers in schools with an Index 1 score between 26 and 50.

Figure 16. Instructional Practice (IP) Rating Distribution by Index 1 Score, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016; TEA Accountability Ratings, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. In 2014–2015, there were no schools with an Index 1 score of less than 25. In 2015–2016, the schools receiving an Index 1 score of less than 25 (n=38) had an Index 1 Target score of 35. Teachers without HR identifying information (n=35) are not included.

- The proportion of teachers with an Index 1 score between 26 and 50 rated as Level 4 increased five percentage points (6% to 11%) and the proportion of teachers rated as Level 3 decreased two percentage points (67% to 65%). At the same time, there was also a decrease in the percentage of teachers in schools with an Index 1 score between 26 and 50 rated Level 2 (23% to 21%) and Level 1 (4% to 3%) (Figure 16).

- Although the majority of teachers at schools with Index 1 scores of 26 to 50 or less were rated Level 3 or Level 4 in 2015–2016 for Instructional Practice, these teachers were twice as likely to be rated Level 1 or Level 2 compared to teachers at schools with Index 1 scores greater than 50 (24% compared to 12%)[6] (Figure 16).

- Similar to the distribution of summative ratings by Index 1 score group, the greatest proportion of teachers with a Level 3 or Level 4 IP rating were located at schools with Index scores of greater than 75 (94%). Though teachers within this group have a lower proportion of teachers rated as Level 3 (56%) compared to other groups of schools by Index 1 score, this group had a higher percentage of teachers

[6] Calculated as the percentage of teachers at schools with index 1 scores greater than 50 who received an Ineffective or Needs Improvement rating (1,242 out of 9,944 teachers).

with Level 4 ratings compared to other groups of schools by Index 1 score. In 2015–2016, the proportion of teachers at schools with Index scores of greater than 75 with rating of Level 4 (38%) surpassed the proportion of teachers with a Level 4 IP rating in schools with Index scores of 51 to 75 (16%) by 22 percentage points (Figure 16, p. 20).

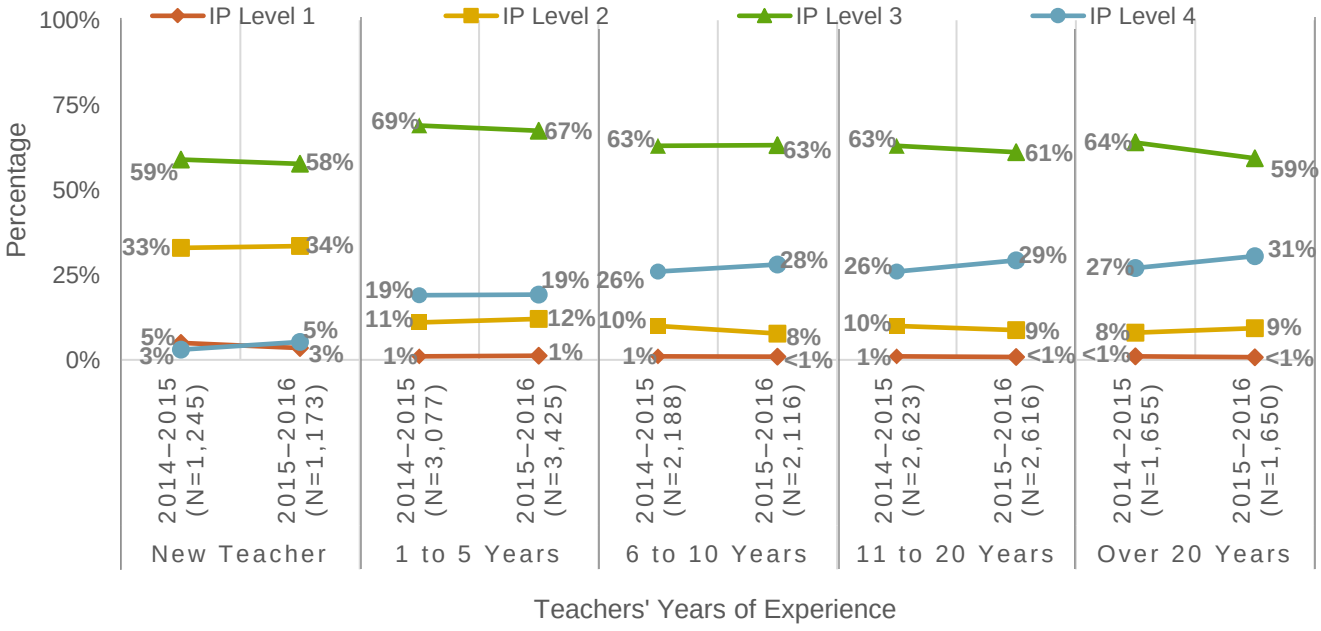
Teachers' Total Years of Experience

- **Figure 17** shows IP ratings by teachers' years of experience for 2014–2015 and 2015–2016. More details on the IP rating distribution by teachers' total years of experience can be found in Appendix E (Table E-2, p. 41). Similar to the summative rating distribution by teachers' years of experience (Figure 10, p.14), the majority of teachers within each group by total years of experience received a Level 3 IP rating. Teachers with 1 to 5 years of experience had the highest proportion of teachers rated as Level 3 (67%) in 2015–2016.

- New teachers, with less than one year of experience, were more than three times more likely to receive a Level 1 or Level 2 IP rating compared to their more experienced colleagues (37% compared to 11% for all other teachers)[7] (Figure 17).

- Consistent with the distribution in summative ratings for teachers in the district from 2014–2015 to 2015–2016, the proportion of teachers with a Level 4 IP rating increased slightly or remained stable for nearly all groups. During the same time, the proportion of teachers receiving an IP rating of Level 3 decreased slightly or remained stable across all groups (Figure 17).

Figure 17. Instructional Practice (IP) Rating Distribution by Teachers' Years of Experience, 2014–2015 and 2015–2016



Sources: Teacher Appraisal and Development F&D Tool, 2014–2015 and 2015–2016; HR Roster File, 2014–2015 and 2015–2016

Note: Percentages may not total 100 due to rounding. Teachers without HR identifying information (n=35) are not included.

[7] Calculated as the percentage of teachers who were not new who received an Ineffective or Needs Improvement rating (1,056 out of 9,807 teachers).

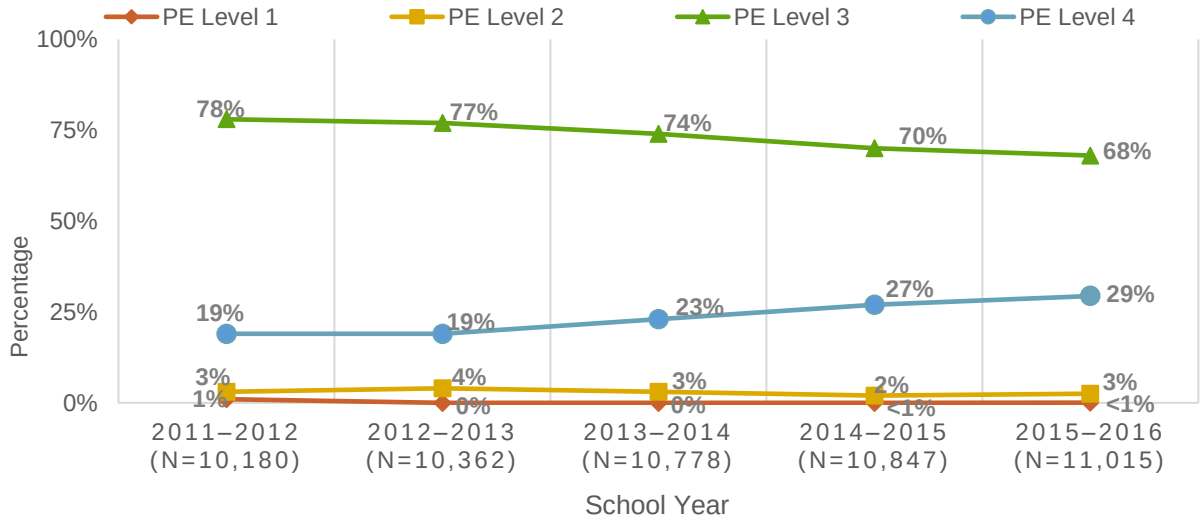
- Teachers with over 20 years of experience saw the greatest change in IP ratings distribution from 2014–2015 to 2015–2016. The proportion of teachers with over 20 years of experience with a Level 4 IP rating increased four percentage points (27% to 31%) and the proportion of teachers with a Level 3 IP rating decreased five percentage points (64% to 59%) (Figure 17, p. 21).

Professional Expectations Ratings

What was the distribution of Professional Expectations (PE) ratings in 2015–2016 compared to previous years?

- **Figure 18** shows the PE ratings distribution from 2011–2012 through 2015–2016. The percentage of teachers with a Level 4 PE rating has steadily increased by ten percentage points, from 19 percent in 2011–2012 and 2012–2013 to 29 percent in 2015–2016. At the same time, the proportion of teachers with a Level 3 PE rating has steadily decreased by ten percentage points, from 78 percent in 2011–2012 to 68 percent in 2015–2016.

Figure 18. Professional Expectation (PE) Rating Distribution 2011–2012 through 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2011–2012, 2012–2013, 2013–2014, 2014–2015 and 2015–2016
Note: Percentages may not equal 100 due to rounding

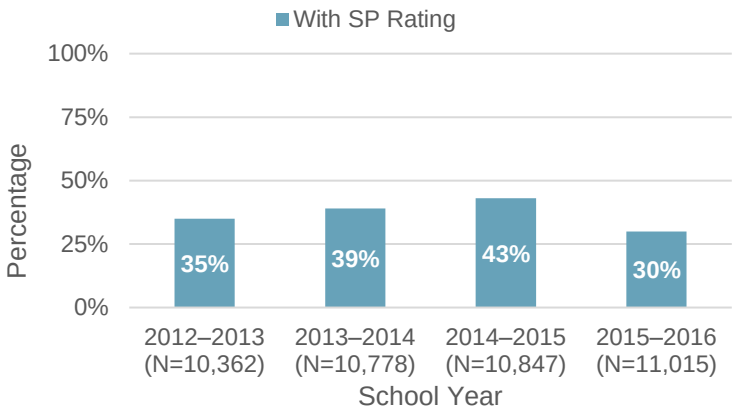
- In 2015–2016, teachers were over four times more likely to receive a Level 1 or Level 2 rating for Instructional Practice than for Professional Expectations (13% compared to 3%) (Figure 13, p. 17, and Figure 18).

Student Performance Ratings

What was the distribution of Student Performance (SP) ratings in 2015–2016 compared to previous years?

- **Figure 19** shows a 13 percentage point decrease in proportion of teachers receiving an SP rating from 2014–2015 to 2015–2016.[8] In 2015–2016, 30 percent (n=3,320) of the 11,015 rated teachers received an SP rating along with an IP and PE rating to make up their overall summative rating.

Figure 19. Teachers with a Student Performance Rating 2012–2013 through 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2012–2013, 2013–2014, 2014–2015, 2015–2016

- **Figure 20** (p. 24) shows the distribution of SP ratings from 2012–2013 to 2015–2016. More details on the SP rating distribution from 2012–2013 to 2015–2016 can be found in Appendix E (**Table E-3**, p. 42). The majority of teachers received an SP rating of Level 4 in 2015–2016. From 2014–2015 to 2015–2016, there was a 32 percentage point increase in the proportion of teachers with a Level 4 rating (35% to 67%). This change coincides with a substantial change in the measures used to calculate SP (see Data Limitations, p. 4).

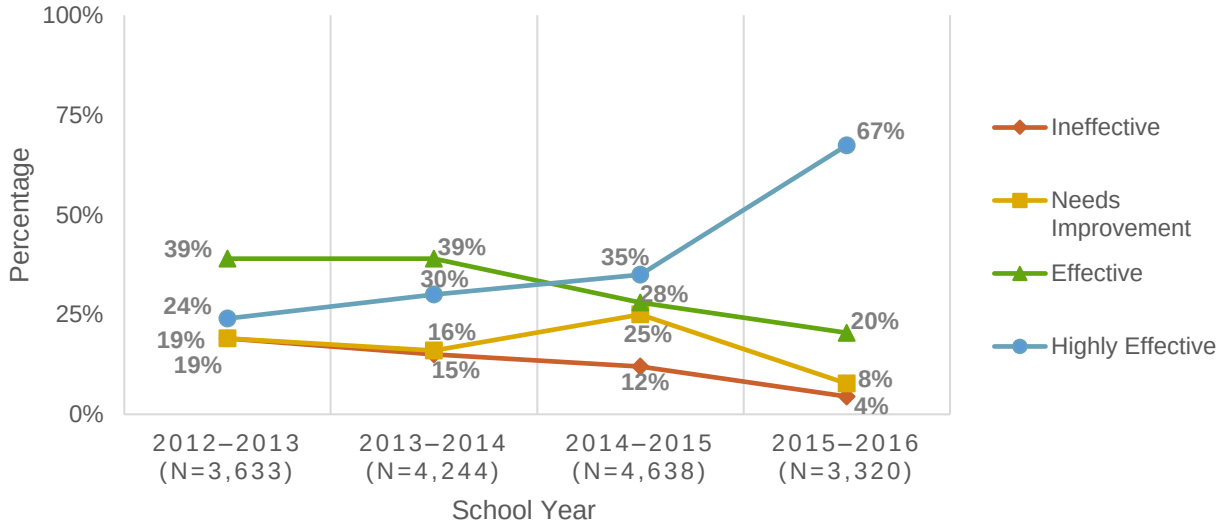
- From 2014–2015 to 2015–2016, there was decrease of eight percentage points in the proportion of teachers rated SP Level 3, from 28 percent to 20 percent. This proportion has decreased steadily since 2012–2013 and 2013–2014, when 39 percent of teachers were rated as Level 3 (Figure 20).

- From 2012–2013 to 2015–2016, there has been variation in the proportion of teachers with a Level 2 SP rating. The proportion of teachers rated as Level 2 was highest in 2014–2015, at 25 percent. From 2014–2015 to 2015–2016, there was a 17 percentage point decrease in the proportion of teachers with a Level 2 rating (25% to 8%) (Figure 20).

[8] The calculation of the Student Performance measures was not consistent between the years. The SP ratings for 2014–2015 included measures of Valued-Added, Comparative Growth for Iowa/Logramos or TELPAS, and Student Progress. The SP ratings for 2015–2016 included measures of Comparative Growth for TELPAS and Student Progress.

- From 2012–2013 to 2015–2016, the proportion of teachers with a Level 1 SP rating has decreased steadily each year to its lowest percentage in 2015–2016 (4%). From 2014–2015 to 2015–2016, there was an eight percentage point decrease in the proportion of teachers with a Level 1 rating (12% to 4%) (Figure 20).

Figure 20. Student Performance (SP) Ratings 2012–2013 through 2015–2016



Source: Teacher Appraisal and Development F&D Tool, 2012–2013, 2013–2014, 2014–2015, 2015–2016
Note: Percentages may not total 100 due to rounding.

- **Table 4** shows the SP rating changes for teachers who received an SP rating in both 2014–2015 and 2015–2016. Of the 3,320 teachers that received an SP rating for 2015–2016, 46 percent of teachers (n=1,541) also received an SP rating for the previous year, 2014–2015.

Table 4. Student Performance (SP) Changes for Teachers with Consecutive SP Ratings, 2014–2015 through 2015–2016 (N=1,541)

	2014–2015 SP Ratings				
2015–2016 SP Ratings	SP Level 1	SP Level 2	SP Level 3	SP Level 4	**Total in 2015–2016**
SP Level 1	13	16	13	5	**47**
SP Level 2	18	34	43	26	**121**
SP Level 3	36	74	119	94	**323**
SP Level 4	34	149	250	617	**1,050**
Total in 2014–2015	**101**	**273**	**425**	**742**	**1,541**

Source: Teacher Appraisal and Development SP Tool, 2014–2015 and 2015–2016
Notes: Percentages may not total 100 due to rounding. Red represents 1) teachers receiving an SP Level 1 rating both years and 2) teachers who fell to SP Level 1 or SP Level 2 in 2015–2016 from a higher rating in 2014–2015. Yellow represents teachers 1) remaining as SP Level 2 both years and 2) teachers who fell from SP Level 4 in 2014–2015 to SP Level 3 in 2015–2016. Green represents 1) teachers who increased their ratings from 2014–2015 to 2015–2016 and 2) teachers remaining as SP Level 3 or SP Level 4.

- A total of 425 teachers who received a Level 3 SP rating in 2014–2015 received an SP rating the next consecutive year. Of those teachers, 59 percent (n=250) increased their SP rating to Level 4 in 2015–-2016 (Table 4, p. 24).

- Of the 273 teachers with a Level 2 SP rating in 2014–2015, 82 percent (n=223) increased their SP rating to Level 3 or Level 4 in 2015–2016 (Table 4).

- Of the 742 teachers who received a Level 4 SP rating in 2014–2015 and received an SP rating the next consecutive year, 83 percent (n=617) maintained a Level 4 SP rating in 2015–2016 (Table 4).

- The mean SP score for the 1,541 teachers with an SP rating for both years increased by 0.26 points from 2014–2015 to 2015–2016. A paired t-test[9] was conducted to compare the mean SP score received by teachers that received an SP rating in both 2014–2015 and 2015–2016. The mean SP score for teachers with an SP rating for both years was significantly higher in 2015–2016 (M = 3.43, SD = 0.77) compared to 2014–2015 (M = 3.17, SD = 0.94), t (1,540) = 10.16, p<0.01.

Student Performance Measures

Given the change in measures available for SP calculations, how did the mean scores for Student Performance measures change in 2015–2016 compared to 2014–2015?

- **Table 5** (p. 26) shows teachers' mean SP scores and mean summative scores[10] by SP measure combination received in 2014–2015 and 2015–2016. More details on the weights used in each component can be found in Appendix A (p. 33). From 2014–2015 to 2015–2016, the proportion of teachers with an SP rating decreased by 13 percentage points (43% to 30%) (Figure 19, p. 23). For only teachers that received an SP rating, the overall mean SP score increased by 0.53 points (2.86 to 3.39) and the mean summative score increased 0.23 points (3.05 to 3.28) from 2014–2015 to 2015–2016. The mean summative score for teachers with an SP rating, 3.28, was 0.12 points higher than the mean summative score for all teachers, 3.16 (Figure 3, p. 6).

- In 2015–2016, the mean SP score and the mean summative score varied by the combination of SP measures, Comparative Growth for TELPAS and Student Progress. There was a 0.55 point difference between the mean SP score for teachers with an SP rating including Comparative Growth, 2.85, and the mean SP score for teachers with an SP rating including Student Progress, 3.40. At the same time, there was a 0.22 point difference between the mean summative score for teachers with an SP rating including Comparative Growth, 3.07, and the mean summative score for teachers with an SP rating including Student Progress, 3.29 (Table 5).

[9] A paired t-test is a statistical procedure to determine whether the difference in means between two groups is significant or due to random chance.
[10] Teachers that receive all three appraisal components (i.e., IP, PE, and SP) received a summative rating based on 50 percent IP, 20 percent PE, and 30 percent SP.

- Between 2014–2015 and 2015–2016, the measures used to calculate Comparative Growth[11] changed and the proportion of teachers with an SP rating that included Comparative Growth decreased 61 percentage points (69% to 8%). At the same time, the mean SP score for teachers with an SP rating including a Comparative Growth measure increased 0.31 points (2.54 to 2.85) (Table 5).

Table 5. Mean Scores by Student Performance Measure Combination, 2014–2015 through 2015–2016

	2014–2015			2015–2016		
Student Performance (SP) Combination	**Number of Teachers**	**Mean SP Score**	**Mean Summative Score**	**Number of Teachers**	**Mean SP Score**	**Mean Summative Score**
Overall SP	4,638	2.86	3.05	3,320	3.39	3.28
SP *including* Value-Added[1]	58% (2,674)	2.53	2.94	-	-	-
SP *including* Comparative Growth[2]	69% (3,207)	2.54	2.93	8% (276)	2.85	3.07
SP *including* Student Progress[3]	42% (1,936)	3.37	3.26	99% (3,293)	3.40	3.29

Source: Teacher Appraisal and Development F&D and SP Tool, 2014–2015 and 2015–2016

Notes: Teachers needed at least two measures within these components to receive an SP rating. See Appendix A (p. 32) for more information.

[1] Indicates all SP measure combinations that include a Value-Added measure. Value-Added was not included as an SP measure for 2015–2016.

[2] Indicates all SP measure combinations that include a Comparative Growth measure. In 2014–2015, CG included Iowa/Logramos or TELPAS assessments. In 2015–2016, CG included TELPAS only.

[3] Indicates all SP measure combinations that include a Student Progress measure.

- From 2014–2015 to 2015–2016, there was a 57 percentage point increase in the proportion of teachers with a SP rating that included Student Progress (42% to 99%). The mean SP score for teachers with an SP rating that included Student Progress increased slightly by 0.03 points (3.37 to 3.40) (Table 5).

- In 2014–2015, teachers with an SP rating including Value-Added (58%) and teachers with an SP rating including Comparative Growth for Iowa/Logramos and/or TELPAS) (69%) received similar mean SP scores (2.53 and 2.54, respectively) and similar mean summative scores (2.94 and 2.93, respectively). (Table 5).

Student Performance and Summative Ratings

What was the impact of Student Performance (SP) on summative ratings in 2015–2016?

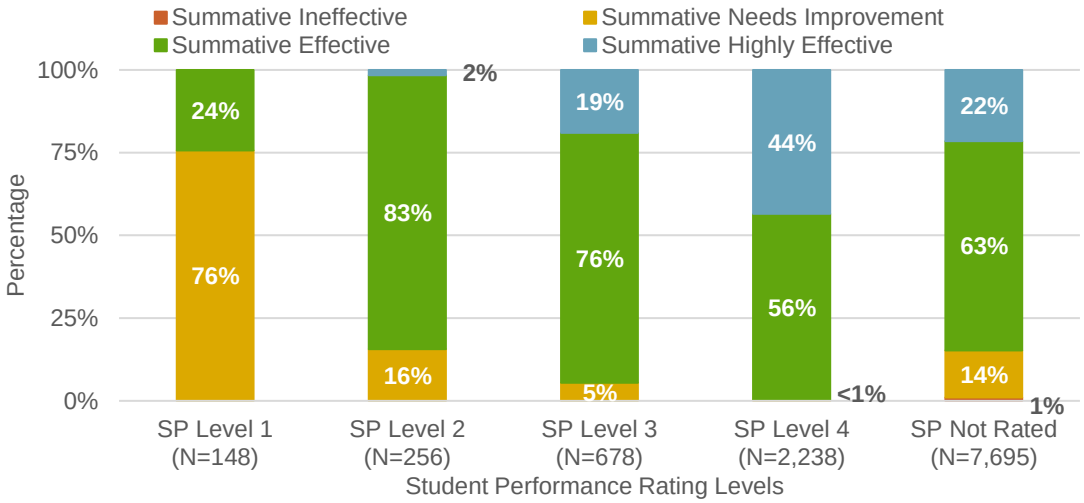
- **Figure 21** (p. 27) shows the 2015–2016 distribution of all summative ratings along each SP performance level, with SP Level 1 indicating teachers with students that did not meet expectations and SP Level 4 indicating teachers with students that exceeded expectations according to the selected student performance measures in 2015–2016. Overall, the data show that summative ratings and SP ratings were generally aligned. In 2015–2016, 95 percent of teachers rated SP Level 3 received a summative rating of Effective or Highly Effective. In addition, nearly 100 percent of teachers with a

[11] For 2014–2015, the Comparative Growth measure included Iowa/Logramos or TELPAS assessments. For 2015–-2016, the Comparative Growth measure included TELPAS only.

Level 4 SP rating received a summative rating of Effective or higher, with the exception of less than one percent (n=6) rated as Needs Improvement.

- Of those teachers with an SP Level 1 rating, 76 percent received a summative rating of Needs Improvement. However, no teachers with an SP Level 1 rating received a corresponding summative rating of Ineffective and 24 percent of them received a summative rating of Effective. Moreover, of teachers with an SP Level 2 rating, 85 percent received a summative rating of Effective or Highly Effective (Figure 21).

Figure 21. Summative Ratings by Student Performance (SP) Levels for All Rated Teachers and Measures, 2015–2016

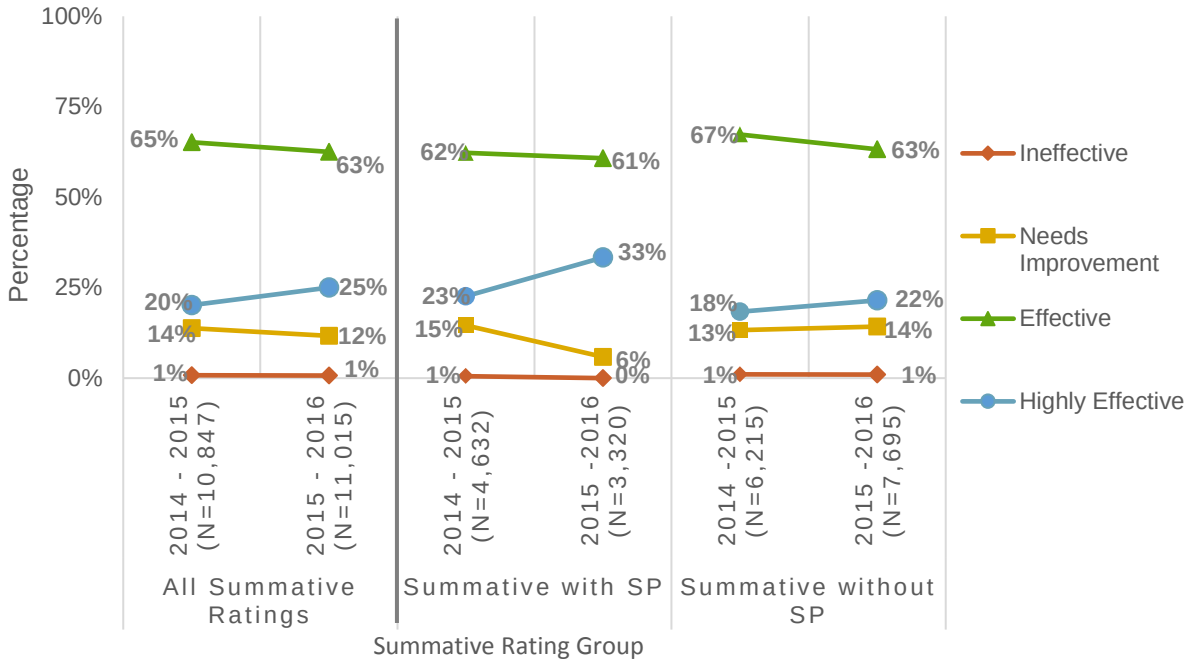


Sources: Teacher Appraisal and Development F&D and SP Tool, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding.

- **Figure 22** (p. 28) shows the summative ratings with an SP rating compared with summative ratings without an SP rating for 2014–2015 and 2015–2016. The proportion of teachers with SP included in their summative rating rated Highly Effective increased by 10 percentage points (23% to 33%) and the proportion of teachers rated as Needs Improvement decreased by nine percentage points (15% to 6%).

- In comparison, the proportion of teachers without SP included in their summative rating rated Highly Effective increased by only four percentage points (18% to 22%) and the proportion of teachers rated as Needs Improvement increased by one percentage point (13% to 14%) (Figure 22).

Figure 22. Summative Ratings with an SP Rating Compared with Summative Ratings without and SP Rating, 2014–2015 and 2015–2016



Source: Teacher Appraisal and Development SP Tool, 2014–2015 and 2015–2016
Note: Percentages may not total 100 due to rounding.

- There was a high variation in the summative ratings distributions for teachers with SP compared to teachers without SP in 2015–2016. The proportion of teachers with a Highly Effective summative rating that included an SP rating (33%) was 11 percentage points higher than for teachers with a Highly Effective summative rating that did not include an SP rating (22%). The proportion of teachers with a Needs Improvement summative rating that included an SP rating (6%) was eight percentage points lower than for teachers with a Needs Improvement summative rating that did not include an SP rating (14%). No teachers with an SP rating received an Ineffective summative rating in 2015–2016 (Figure 22).

Discussion

The 2015–2016 school year marked the fifth year of TADS as HISD's teacher appraisal system. From 2014–2015 to 2015–2016, the percentage of teachers appraised in the TADS system decreased by four percentage points (94% to 90%). Analysis of summative rating trends in 2015–2016 indicates the proportion of Effective or Highly Effective-rated teachers was the highest (88%) since TADS was introduced in 2011–2012, including a five percentage point increase (20% to 25%) in the proportion of teachers rated as Highly Effective.

When separated by individual TADS components, both Instructional Practice (IP) and Professional Expectations (PE) Level 4 ratings, which are scored based on appraisers' observations, increased from 2011–2012 to 2015–2016. One explanation for this incremental increase in IP and PE could be that the TADS system was proficient in identifying teachers' areas of instructional growth and facilitating targeted support. In other words, quality, individualized feedback from appraisers may be providing teachers with information that improves their performance in the classroom. Research also suggests that there may be

unintentional reasons for increases to the observational components of teacher evaluation systems. One report found that teachers, regardless of true performance, tend to receive high ratings in evaluation systems. Known as the Widget Effect, this pattern was attributed to evaluation systems' lack of differentiation for the variations of teacher effectiveness (Weisberg, et al., 2009). Another possibility for increases in IP and PE over the five years could be related to the teacher-appraiser relationship (e.g., quality of appraiser recommendations for support, appraiser bias towards a highly regarded teacher, etc.) (Education Analytics, 2016). HISD might consider studying possible correlations between TADS and the facilitation of instructional support to improve teacher effectiveness, possibly by analyzing longitudinal and qualitative data for teachers and appraisers that have participated in TADS across multiple years.

The distribution of Student Performance (SP) ratings were not consistent with IP and PE distribution rating trends in 2015–2016. The differences may be, at least partially, attributed to Student Performance measures changing from 2014–2015 to 2015–2016. The SP ratings for 2014–2015 included measures of Value-Added, Comparative Growth for Iowa/Logramos (i.e., normative assessments) or TELPAS, and Student Progress. By board decision, the district excluded Value-Added and normative assessments from the SP calculations the following year, 2015–2016. As a result, SP ratings for 2015–2016 included only measures of Comparative Growth for TELPAS and Student Progress, a subjective, less-rigorous measure. In 2015–2016, teachers were twice as likely to receive a rating of Level 4 for SP (67%) compared to IP (24% of teachers were rated Level 4) and PE (29% of teachers were rated Level 4).

From 2014–2015 to 2015–2016, the proportion of teachers with an SP rating included in their summative appraisal rating fell 13 percentage points (43% to 30%). At the same time, the proportion of teachers with SP included in their summative rating rated as Highly Effective increased by 10 percentage points (23% to 33%) and the proportion of teachers rated as Needs Improvement decreased by nine percentage points (15% to 6%). For those teachers that received an SP rating, the mean SP score increased by 0.53 points (2.86 to 3.39) and the mean summative score increased 0.23 points (3.05 to 3.28) from 2014–2015 to 2015–2016. Changes to the Student Performance measurements from year to year may impact the proportion of teachers able to incorporate SP into their summative rating, the score for the SP rating, and consequently, the overall summative score. It is recommended that the district consider maintaining the same SP measures for at least two consecutive years to assess the impact of SP measure changes on teachers' ability to use SP in their summative rating and SP rating score.

Student Progress, similar to student learning objectives (SLOs), is a participatory student growth measure in which teachers and their appraisers work together to set appropriate goals for students. Recent studies call for more rigorous research on the validity of using student learning goals to measure the relationship between student achievement and teacher effectiveness (Tyler, 2011; Lacireno-Paquet, Morgan, & Mello, 2014). Further analysis of the Student Performance component in 2015–2016 indicates variations in the mean SP and summative scores by SP measures. Specifically, when a Student Progress measure is included in the SP rating, teachers, on average, appear to receive a higher SP rating than when Student Progress is not included in SP. From 2014–2015 to 2015–2016, the proportion of teachers with Student Progress included in their SP rating increased by 57 percentage points (42% to 99%). For both 2014–2015 and 2015–2016, the mean SP score and the mean summative score were higher for teachers that included Student Progress in their SP rating than for teachers who did not include Student Progress. In 2015–2016, specifically, teachers that included Student Progress as an SP measure, on average, received an SP rating 0.55 points higher than teachers that included Comparative Growth as an SP measure (3.40 compared to 2.85, respectively). At the same time, there was a 0.23 point difference between the mean summative score for teachers with an SP rating including Comparative Growth, 3.07, and the mean summative score for

teachers with an SP rating including Student Progress, 3.29. These findings suggest that the Student Progress measure may increase a teacher's SP score, and consequently, the overall summative rating.

Findings from trends of summative rating distributions by groups on the campus- and teacher-levels offer insight for the district on possibilities for targeted professional development or TADS system improvements. Since 2012–2013, Met Standard schools have had higher proportions of teachers with a summative rating of Highly Effective than IR schools. In 2015–2016, Met Standard schools had more than double the proportion of Highly Effective teachers (27%) compared to teachers at IR schools (12%). At the same time, teachers in IR schools (23%) were more than twice as likely to be rated as Needs Improvement compared to teachers in Met Standard schools (10%). HISD should focus on growing and supporting teachers at campus and teacher level groups who have consistently received lower ratings than teachers in comparison groups. In addition, HISD should continue efforts that attract effective teachers to IR schools. Future research could explore the possible reasons why teachers at IR schools received lower ratings through analysis of longitudinal and qualitative data for teachers and appraisers that have participated in TADS across multiple years.

In 2015–2016, the proportion of teachers with one to five years of experience with a summative rating of Highly Effective was nine percentage points lower than teachers with six to 10 years of experience. At the same time, the proportion of teachers with one to five years of experience with an Effective summative rating was five percentage points higher than teachers with six to 10 years of experience. This finding may be explained by research from Henry, Bastian, & Fortner (2011), which found that teachers who continue teaching after their fifth year are, on average, more effective than teachers who exit after their third or fourth year of teaching. The study notes that teachers continue to make substantial gains in effectiveness through their third year of teaching when provided with appropriate coaching and professional development (Henry, Bastian, & Fortner, 2011). In other words, it is possible that the different proportions of Highly Effective and Effective ratings between teachers with less than five years of experience and teachers with six to 10 years of experience is because less effective teachers have been leaving the teaching profession before their fifth year of teaching. However, support for teachers after their first year of teaching with less than five years of experience may lead to considerable gains in effective instructional practice. HISD should continue to support ongoing strategies for these teachers, including efforts to retain effective teachers beyond their first few years of teaching.

This report has examined teacher appraisal outcomes for the 2015–2016 school year and previous years. Trends observed in appraisal outcomes by campus- and teacher-level groups can guide decision-makers in their work toward accurately rating effective teaching, strengthening professional development and support, growing teachers' capacity for effective teaching, and ultimately placing an effective teacher in every classroom.

References

Education Analytics, Inc. (2016). Teachers Appraisal and Development System Research Study. Madison, WI: Education Analytics, Inc.

Hedges, L. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics*. Vol. 6, No. 2 (Summer, 1981), pp. 107-128.

Henry, G. T., Bastian, K. C., & Fortner, C. K. (2011). Stayers and leavers: Early-career teacher effectiveness and attrition. *Educational Researcher*, *40*(6), 271-280.

Lacireno-Paquet, N., Morgan, C., & Mello, D. (2014). How states use student learning objectives in teacher evaluation systems: a review of state websites (REL 2014–013). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov /ncee/edlabs.

Leadership Development. (2013). HISD Teacher Appraisal and Development System Manual, 2013–2014. Houston, TX: Houston Independent School District.

Leadership Development. (2015a). HISD Teacher Appraisal and Development System: Instructional Practice and Professional Expectations Rubrics. Houston, TX: Houston Independent School District.

Leadership Development. (2015b). HISD Teacher Appraisal and Development System: Student Performance Guidebook. Houston, TX: Houston Independent School District.

Research and Accountability. (2015). Teacher Appraisal and Development System End of Year Report, 2013–2014. Houston, TX: Houston Independent School District.

Research and Accountability. (2016). Teacher Appraisal and Development System End of Year Report, 2014–2015. Houston, TX: Houston Independent School District.

Texas Education Agency. (2016). State Accountability Manual, 2016. Austin, TX: Texas Education Agency. Retrieved from http://tea.texas.gov/2016accountabilitymanual.aspx

Texas Education Agency. (2017). Texas Equity Toolkit. Austin, TX: Texas Education Agency. Retrieved from https://texasequitytoolkit.org/

Tyler, J. H. (2011). Designing High Quality Evaluation Systems for High School Teachers: Challenges and Potential Solutions. Washington, DC: Center for American Progress.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness* (2nd ed.). New York, NY: New Teacher Project. Retrieved from https://tntp.org/assets/documents/TheWidgetEffect_execsummary_2nd_ed.pdf

Whitehurst, G. J. R., Chingos, M. M., & Lindquist, K. M. (2014). *Evaluating teachers with classroom observations.* Washington, DC: Brown Center on Education Policy. Retrieved from https://www.brookings.edu/research/evaluating-teachers-with-classroom-observations-lessons-learned-in-four-districts/

APPENDIX A: TADS Components Distribution

TADS Components Distribution 2015 – 2016 *ONLY*

The component weights are applied to derive the Summative Appraisal Rating (IP, PE, and SP combined).[1]

Ineffective 1.00 – 1.49	Needs Improvement 1.50 – 2.49	Effective 2.50 – 3.49	Highly Effective 3.50 – 4.00

Teachers with <u>two TADS components</u> (i.e. no Student Performance rating) have the following weights within teachers' Summative Appraisal Ratings.

Instructional Practice (IP)	Professional Expectations (PE)
70%	30%

Teachers with <u>three TADS components</u> have the following weights within teachers' Summative Appraisal Ratings.

Instructional Practice (IP)	Professional Expectations (PE)	Student Performance (SP)
50%	20%	30%

The various types of <u>Student Performance[2] [3]</u> measures have different weights within the Student Performance final rating.

SP Measure / SP Measure Combinations	Comparative Growth (CG)[4]	Student Progress[5]	Value-Added	Student Performance (SP) Total
CG Only	30%			30%
CG + Student Progress	20%	10%		30%
Student Progress Only		30%		30%
CG + Value-Added +	N/A		N/A	N/A
Student Progress + Value-Added		N/A	N/A	N/A
CG + Student Progress + Value-Added	N/A	N/A	N/A	N/A

[1] All TADS components, including Student Performance (SP) measures of Comparative Growth and Student Progress, use a 4-point scale
[2] Teachers must have a minimum of two Student Performance measures to receive a Student Performance rating included in the summative rating.
[3] On June 9, 2016, the HISD Board voted not to extend the contract with SAS EVAAS®. Teacher-level Value-Added Growth (EVAAS®) for the 2015–2016 TADS is not included in the Student Performance or Summative Ratings.
[4] CG is a district measure based on TELPAS and/or STAAR assessments in certain grade levels and subjects.
[5] Student Progress is a student learning measure that uses *two measures of* a) district-wide/pre-approved/appraiser-approved assessments, b) district-wide/pre-approved/appraiser-approved performance tasks/work products, or c) student attainment (Pre-K teachers only).

Last Updated 4/18/2017

Source: Leadership Development, 2015b

APPENDIX B: Guide to the TADS Summative Component Distribution

HISD Teacher Appraisal and Development System			
Measure		**Summative Rating Weight**	**Criteria for Measurement**
Instructional Practice Criteria	Planning (PL)	50% OR 70%	PL-1 Develops student learning goals
			PL-2 Collects, tracks, and uses student data to drive instruction;
			PL-3 Designs effective lesson plans, units, and assessments
	Instruction (I)		I -1 Facilitates organized, student-centered, objective-driven lessons
			I-2 Checks for student understanding and responds to student misunderstanding
			I-3 Differentiates instruction for student needs by employing a variety of instructional strategies
			I-4 Engages students in work that develops higher-level thinking skills
			1-5 Maximizes instructional time
			1-6 Communicates content and concepts to students
			1-7 Promotes high expectations for students
			1-8 Students actively participating in lesson activities
			1-9 Sets and implements discipline management procedures
			1-10 Builds a positive and respectful classroom environment
Professional Expectations Criteria	Professionalism (PR)	20% OR 30%	PR-1 Complies with policies and procedures at school
			PR-2 Treats colleagues with respect throughout all aspects of work
			PR-3 Complies with teacher attendance policies
			PR-4 Dresses professionally according to school policy
			PR-5 Collaborates with colleagues
			PR-6 Implements school rules
			PR-7 Communicates with parents throughout the year
			PR-8 Seeks feedback in order to improve performance
			PR-9 Participates in professional development and applies learning
Student Performance Criteria	Student Performance (PR)	30% OR N/A	*Value-Added not included in 2015 – 2016 summative rating*
			Comparative Growth (CG) on TELPAS grades 3–8
			Student Progress • On districtwide, pre-approved, or appraiser-approved assessments • On districtwide, pre-approved, or appraiser-approved tasks • Student attainment (Pre-K only)

Source: Leadership Development, 2013; Leadership Development, 2015b

APPENDIX C: 2015–2016 Student Performance Measures in Detail

The ***Student Performance Rating (SP***) is a composite metric used in teachers' appraisal ratings when applicable. Teachers must have at least two of the following measures for SP to be applied to their overall summative rating:

- Comparative Growth on district-wide assessments;
- Students' progress on districtwide assessments, pre-approved assessments, or appraiser-approved assessments
- Students' progress on districtwide, pre-approved, or appraiser-approved performance tasks or products
- Student attainment on districtwide or appraiser-approved assessments.

SP ratings are on a scale of 1–4. A teacher must have at least two SP measures to receive an SP rating. Teachers who do not receive an SP rating will receive a Summative Appraisal Rating based solely on an Instructional Practice (IP) rating and a Professional Expectations (PE) rating assigned by the appraiser.

Measure #1: Value-Added Growth

Value-Added Growth is a district-rated measure of the extent to which a student's average growth meets, exceeds, or falls short of average growth of students in the district. Value-added analysis uses a student's own academic performance across years, grades, and subjects as a basis for determining his/her average growth. EVAAS® was used as the value-added growth measure for teachers with available data in the Student Performance (SP) rating for TADS from 2011–2012 through 2014–2015.

On June 9, 2016, the HISD Board voted not to extend the contract with SAS EVAAS®. As a result, teacher-level Value-Added Growth for the 2015–2016 Teacher Appraisal and Development System (TADS) is not included in the Student Performance or Summative Ratings. If teachers do not have two Student Performance measures besides Value-Added, their summative rating will be calculated using only their Instructional Practice and Professional Expectations ratings.

Measure #2: Comparative Growth on district-wide assessments

Comparative growth may be used to calculate a teacher's Student Performance rating. Comparative Growth (CG) measures the progress of a teacher's students on a given assessment compared to all other students within the same school district who start at the same test-score level. For 2015–2016, CG relies on the use of TELPAS assessments in grades 3–8, and is computed using two consecutive years of students' scores. CG scores are placed on a scale of 1–4. For the 2015 -2016 school year, only TELPAS assessments for Reading scale scores in grades 3–8 were used to calculate the CG measure. In prior years, CG was calculated using norm-referenced data in grades 2–8.

Measure #3 & #4: Student Progress

Student Progress may be used to calculate a teacher's Student Performance rating. Student Progress is a type of TADS Student Performance measure that uses summative or cumulative assessments, performance tasks, and work products to measure how much content and skill students learned over the duration of a course or year, based on where they started the subject or course.

#3. Students' progress on districtwide assessments, pre-approved assessments, or appraiser-approved assessments

Student Progress is a student learning measure that uses summative assessments to measure how much content and skill students learned over the duration of a course or year, based on where they started the subject or course. Student Progress is an appraiser-approved rating of the extent to which students learned an ambitious and feasible amount of content and skills, taking into account students' starting points. To measure Student Progress, teachers must create Goals Worksheets for no more than two of the courses they teach and place students into appropriate starting points based on two pieces of evidence, such as past grades or past test scores. Once students have been placed into an appropriate starting group, which must be approved by the teacher's appraiser, they will receive a goal dependent upon which assessment is appropriate for that course. Assessment results are entered into a Results Worksheet either automatically or by the teacher. Once the Results Worksheets have been approved by the appraiser, a teacher will receive a Performance Level rating based on how many students achieved their goals. Performance Levels are on a scale of 1–4.

#4. Students' progress on districtwide, pre-approved, or appraiser-approved performance tasks or products

The Student Progress process using appraiser-approved culminating performance tasks or work products mirrors the process for Student Progress on assessments. The only substantive difference is the type of summative assessment tool used. For example, in certain subjects, such as art, music, or foreign language, a culminating project or performance task might be more appropriate than, or used in conjunction with, a more traditional paper-pencil test.

Measure #5: Student Attainment

Student Attainment is a student learning measure that uses district-wide or appraiser-approved assessments to measure how many students performed at a target level, regardless of their starting point. Currently, Student Attainment only applies to Pre-K.

Source: Leadership Development, 2015b, pp. 35–40

APPENDIX D: Methodology for TADS End of Year Report, 2015–2016

For this report, HISD Human Resources (HR) provided district-wide employee rosters, which included multiple identifiers for teacher-level data. The ASPIRE (Accelerating Student Progress: Increasing Results and Expectations) team, housed within the HISD Department of Research and Accountability, provided additional identifiers for teacher-level data. Additional data on school accountability was obtained through the Texas Education Agency (TEA) website. A teacher was eligible for appraisal if s/he was actively employed from the beginning of the school year until the end of April of each academic year. In each case, only teachers who received a TADS summative rating were included in the analyses.

- Teachers were identified using the following criteria:
    - To identify job descriptions specific to teachers, the variable *Job Function Code* was reported as TCH, TEA ELEM, TEA PREK, TEA SEC, or # (i.e., not assigned job function code)
    - To identify salary plans specific to teachers, the variable *Salary Plan* was reported as RT, VT, RO1 or RO5.

- A teacher's school level was determined by identifying campus level assignments specific to each teacher as indicated in the 2015–2016 HISD District and School Profiles. Teachers located at Community Services, HCC Life Skills, EL DAEP, and Beechnut Academy were excluded in this report.

- Campus accountability ratings and Index 1 Scores were obtained from the TEA using the Texas Academic Performance Reports (TAPR) for 2012–2013, 2013–2014, 2014–2015, and 2015–2016. TEA did not release accountability data for 2011–2012.

- Teachers coded by ASPIRE in the student-linkage database as Core Foundation teachers included those who taught courses in math, science, social studies, English, and reading.

- Critical shortage teachers were identified as teaching in a TEA-defined critical shortage or high needs area. To be included in this category, the variable *Job Family* was reported as BIL, MATH, SCIENC, SPECIAL ED, ESL, CATE, and/or COMP.

- Total retention was defined as those teachers from the 2015–2016 school year who remained actively employed in HISD in 2016–2017, including those no longer assigned to classrooms. Teachers retained in the district were reported by HISD Human Resources Information System (HRIS) as receiving a status code of A (active), B (paid leave), C (unpaid leave), F (FMLA Full), or E (FMLA Int). Teachers were considered as retained if they were employed from May 2016 through August 2016.

- Teacher movement was defined as teachers who stayed in the district (those retained) who changed locations within HISD from May 2016 to August 2016, regardless of whether the location change included a promotion.

- Teacher years of experience was determined using total professional experience as verified by HRIS. Total professional experience is defined as the sum of the variables, *HISD Experience* and *Other Experience*. Teachers were categorized as new teachers (i.e., in their first year of teaching), 1–5 years, 6–10 years, 11–20 years, and more than 20 years of experience.

- Comparative Growth (CG) is one of five possible measures for Student Performance. CG looks at the progress of a teacher's students compared to all other students within the same school district who start at the same level. This measure is calculated by the ASPIRE team, housed within the HISD Department of Research and Accountability. For 2015–2016, CG was a district measure based on Texas English Language Proficiency Assessment System (TELPAS) assessments for grades 3–8. Data were loaded into the TADS F&D Tool for appraisal purposes.

Appendix E: Data Tables

Table E-1. 2011–2012 through 2015–2016 Summative Rating Distribution by Campus and Teacher Characteristics

	Ineffective (N)					Needs Improvement (N)					Effective (N)					Highly Effective (N)					Totals (N)				
	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016
Overall Summative Rating	1% (109)	3% (313)	3% (326)	1% (91)	1% (78)	12% (1,227)	19% (2,001)	17% (1,799)	14% (1,500)	12% (1,289)	61% (6,235)	59% (6,125)	59% (6,334)	65% (7,067)	63% (6,886)	26% (2,609)	19% (1,923)	22% (2,319)	20% (2,189)	25% (2,762)	100% (10,180)	100% (10,362)	100% (10,778)	100% (10,847)	100% (11,015)
School Level																									
Elementary	1% (54)	4% (213)	3% (180)	1% (54)	1% (42)	12% (712)	23% (1,285)	18% (1,079)	15% (897)	12% (737)	61% (3,473)	57% (3,250)	58% (3,441)	65% (3,862)	63% (3,808)	26% (1,460)	17% (957)	21% (1,256)	19% (1,134)	24% (1,469)	100% (5,699)	100% (5,705)	100% (5,956)	100% (5,947)	100% (6,056)
Middle	1% (24)	4% (67)	5% (91)	1% (14)	1% (18)	12% (206)	20% (346)	20% (361)	18% (319)	14% (264)	62% (1,046)	54% (937)	55% (1,005)	65% (1,182)	66% (1,209)	25% (422)	22% (387)	20% (359)	16% (295)	18% (331)	100% (1,698)	100% (1,737)	100% (1,816)	100% (1,810)	100% (1,822)
High	1% (22)	1% (23)	2% (42)	1% (16)	1% (15)	11% (247)	12% (271)	12% (280)	9% (213)	9% (221)	62% (1,398)	69% (1,587)	67% (1,551)	67% (1,559)	63% (1,481)	26% (586)	19% (430)	19% (453)	23% (524)	28% (652)	100% (2,253)	100% (2,311)	100% (2,326)	100% (2,312)	100% (2,369)
Combined	2% (9)	2% (10)	2% (13)	<1% (5)	<1% (3)	12% (62)	16% (99)	12% (79)	9% (60)	10% (66)	60% (318)	58% (351)	50% (332)	58% (396)	47% (325)	27% (141)	24% (149)	37% (246)	32% (219)	43% (296)	100% (530)	100% (609)	100% (670)	100% (680)	100% (690)
Total (N)	**109**	**313**	**326**	**89**	**78**	**1,227**	**2,001**	**1,799**	**1,489**	**1,288**	**6,235**	**6,125**	**6,329**	**6,999**	**6,823**	**2,609**	**1,923**	**2,314**	**2,172**	**2,748**	**10,180**	**10,362**	**10,768**	**10,749** [2]	**10,937** [5]
Accountability Rating†																									
Improvement Required (IR)	-	5% (89)	6% (101)	2% (37)	1% (21)	-	29% (530)	28% (478)	27% (521)	23% (318)	-	57% (1,027)	57% (968)	65% (1,253)	64% (900)	-	9% (156)	9% (153)	6% (121)	12% (166)	-	100% (1,802)	100% (1,700)	100% (1,932)	100% (1,405)
Met Standard	-	3% (224)	2% (225)	1% (52)	1% (57)	-	17% (1471)	15% (1,320)	11% (967)	10% (966)	-	60% (5,098)	59% (5,357)	65% (5,739)	62% (5,915)	-	21% (1,767)	24% (2,161)	23% (2,049)	27% (2,580)	-	100% (8,560)	100% (9,063)	100% (8,807)	100% (9,518)
Total (N)	**-**	**313**	**326**	**89**	**78**	**-**	**2,001**	**1,798**	**1,488**	**1,284**	**-**	**6,125**	**6,325**	**6,992**	**6,815**	**-**	**1,923**	**2,314**	**2,170**	**2,746**	**-**	**10,362**	**10,763**	**10,739** [3]	**10,923** [6]
Index 1 Scores†																									
25 or Less	-	5% (3)	25% (3)	0% (0)	0% (0)	-	24% (13)	25% (3)	0% (0)	26% (10)	-	67% (37)	42% (5)	0% (0)	61% (23)	-	4% (2)	8% (1)	0% (0)	13% (5)	-	100% (55)	100% (12)	100% (0)	100% (38)
26 to 50	-	6% (24)	10% (68)	3% (36)	1% (14)	-	32% (128)	35% (246)	28% (359)	21% (207)	-	53% (214)	52% (365)	65% (840)	66% (661)	-	9% (38)	3% (24)	4% (57)	12% (116)	-	100% (404)	100% (703)	100% (1,292)	100% (998)
51 to 75	-	4% (232)	4% (203)	1% (50)	1% (54)	-	23% (1,247)	21% (1,134)	16% (957)	15% (885)	-	61% (3,290)	62% (3,327)	69% (4,133)	67% (4,055)	-	11% (597)	14% (742)	14% (852)	18% (1,076)	-	100% (5,366)	100% (5,406)	100% (5,992)	100% (6,070)
Greater than 75	-	1% (54)	1% (52)	<1% (3)	<1% (10)	-	13% (609)	9% (413)	5% (173)	5% (186)	-	57% (2,566)	57% (2,615)	59% (2,062)	55% (2,123)	-	28% (1,285)	33% (1,547)	36% (1,266)	40% (1,555)	-	100% (4,514)	100% (4,627)	100% (3,504)	100% (3,874)
Total (N)	**-**	**313**	**326**	**89**	**78**	**-**	**1,997**	**1,796**	**1,488**	**1,288**	**-**	**6,107**	**6,312**	**6,992**	**6,862**	**-**	**1,922**	**2,314**	**2,170**	**2,752**	**-**	**10,339**	**10,748**	**10,788** [4]	**10,980** [7]
Core Foundation Teachers																									
Core	1% (75)	4% (274)	4% (275)	1% (68)	1% (57)	13% (871)	22% (1,556)	19% (1,403)	15% (1,238)	12% (1,064)	60% (4,151)	55% (3,829)	55% (4,013)	65% (5,447)	62% (5,378)	26% (1,773)	19% (1,320)	22% (1,564)	20% (1,657)	25% (2,146)	100% (6,870)	100% (6979)	100% (7,255)	100% (8,410)	100% (8,645)
Non-Core	1% (34)	1% (39)	1% (51)	1% (21)	1% (21)	11% (356)	13% (445)	11% (396)	11% (251)	10% (224)	63% (2,084)	68% (2,296)	66% (2,318)	67% (1,588)	64% (1,484)	25% (836)	18% (603)	21% (755)	22% (518)	26% (606)	100% (3,310)	100% (3,383)	100% (3,520)	100% (2,378)	100% (2,335)
Total (N)	**109**	**313**	**326**	**89**	**78**	**1,227**	**2,001**	**1,799**	**1,489**	**1,288**	**6,235**	**6,125**	**6,331**	**7,035**	**6,862**	**2,609**	**1,923**	**2,319**	**2,175**	**2,752**	**10,180**	**10,362**	**10,775**	**10,788** [4]	**10,980** [7]

Source: TADS Feedback and Development Tool; TADS Student Performance Tool; HISD PeopleSoft Rosters: 2011–2012 as of 04–16–2012; 2012–2013 as of 04–10–2013; 2013–2014 as of 04–14–2014; 2014–2015 as of 05–15–2015; 2015–2016 as of 05–28–2016

†Accountability ratings not available for school year 2011–2012.

* Retention and Teacher Movement were not calculated in the TADS End of Year Reports for 2011–2012; 2012–2013; and 2013–2014

[1] 36 teachers excluded from data.

[2] 98 teachers at Community Services, HCC Life Skills, EL DAEP, or with no school identifying information in HR Roster. Not included in school levels.

[3] 108 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[4] 59 teachers without HR Roster identifying information.

[5] 78 teachers at Community Services, HCC Life Skills, El DAEP, Beechnut Academy, or with no school identifying information in HR Roster. Not included in school levels.

[6] 57 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[7] 35 teachers without HR Roster identifying information.

Table E-1 *continued*

	Ineffective (N)					Needs Improvement (N)					Effective (N)					Highly Effective (N)					Totals (N)				
	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016
Critical Shortage Teachers																									
Critical Shortage	1% (34)	3% (72)	3% (75)	1% (30)	1% (26)	13% (313)	18% (442)	16% (408)	13% (518)	12% (573)	63% (1,517)	62% (1,533)	63% (1,556)	67% (2,744)	62% (3,063)	22% (525)	17% (414)	18% (450)	19% (789)	26% (1,277)	100% (2,389)	100% (2,461)	100% (2,489)	100% (4,081)	100% (4,939)
Non-Critical Shortage	1% (75)	3% (241)	3% (251)	1% (59)	1% (52)	12% (914)	20% (1,559)	17% (1,391)	14% (971)	12% (715)	61% (4,718)	58% (4,592)	58% (4,775)	64% (4,291)	63% (3,799)	27% (2,084)	19% (1,509)	23% (1,869)	21% (1,386)	24% (1,475)	100% (7,791)	100% (7,901)	100% (8,286)	100% (6,707)	100% (6,041)
Total (N)	**109**	**313**	**326**	**89**	**78**	**1,227**	**2,001**	**1,799**	**1,489**	**1,288**	**6,235**	**6,125**	**6,331**	**7,035**	**6,862**	**2,609**	**1,923**	**2,319**	**2,175**	**2,752**	**10,180**	**10,362**	**10,775**	**10,788** [4]	**10,980** [7]
Years of Experience																									
New Teacher	2% (16)	7% (102)	8% (149)	3% (32)	2% (23)	28% (211)	33% (510)	30% (595)	35% (430)	32% (377)	59% (444)	54% (843)	54% (1,061)	59% (740)	60% (704)	11% (80)	7% (109)	8% (165)	3% (43)	6% (69)	100% (751)	100% (1,564)	100% (1,970)	100% (1,245)	100% (1,173)
1-5 Years	1% (26)	2% (64)	3% (89)	1% (24)	1% (27)	10% (361)	17% (476)	14% (464)	12% (372)	12% (398)	64% (2,213)	60% (1,721)	61% (1,983)	69% (2,136)	67% (2,281)	25% (872)	22% (622)	22% (723)	18% (545)	21% (719)	100% (3,472)	100% (2,883)	100% (3,259)	100% (3,077)	100% (3,425)
6-10 Years	1% (24)	2% (47)	1% (31)	1% (13)	<1% (10)	11% (259)	17% (382)	13% (277)	12% (261)	8% (161)	62% (1,459)	62% (1,407)	60% (1,258)	64% (1,406)	62% (1,318)	26% (641)	19% (441)	26% (541)	23% (508)	30% (627)	100% (2,419)	100% (2,277)	100% (2,107)	100% (2,188)	100% (2,116)
11-20 Years	1% (22)	3% (59)	1% (29)	<1% (13)	<1% (12)	12% (261)	17% (387)	14% (303)	11% (286)	8% (204)	58% (1,288)	59% (1,338)	58% (1,291)	64% (1,682)	62% (1,611)	29% (634)	21% (483)	27% (610)	24% (642)	30% (789)	100% (2,205)	100% (2,267)	100% (2,233)	100% (2,623)	100% (2,616)
Over 20 Years	2% (21)	3% (35)	2% (27)	<1% (7)	<1% (6)	10% (135)	18% (224)	13% (158)	8% (140)	9% (148)	60% (795)	59% (736)	61% (737)	65% (1,071)	57% (948)	29% (382)	21% (260)	23% (280)	26% (437)	33% (548)	100% (1,333)	100% (1,255)	100% (1,202)	100% (1,655)	100% (1,650)
Total (N)	**109**	**307**	**325**	**89**	**78**	**1,227**	**1,979**	**1,797**	**1,489**	**1,288**	**6,199**	**6,045**	**6,330**	**7,035**	**6,862**	**2,609**	**1,915**	**2,319**	**2,175**	**2,752**	**10,144** [1]	**10,246**	**10,771**	**10,788** [4]	**10,980** [7]
Retention*																									
Retained	-	-	-	<1% (28)	<1% (28)	-	-	-	12% (1,080)	10% (919)	-	-	-	67% (6,189)	64% (6,056)	-	-	-	21% (1,969)	26% (2,508)	-	-	-	100% (9,266)	100% (9,511)
Exited	-	-	-	4% (61)	3% (50)	-	-	-	27% (409)	25% (369)	-	-	-	56% (846)	55% (806)	-	-	-	13% (206)	17% (244)	-	-	-	100% (1,522)	100% (1,469)
Total (N)	**-**	**-**	**-**	**89**	**78**	**-**	**-**	**-**	**1,489**	**1,288**	**-**	**-**	**-**	**7,035**	**6,862**	**-**	**-**	**-**	**2,175**	**2,752**	**-**	**-**	**-**	**10,788** [4]	**10,980** [7]
Teacher Movement*																									
Remained at the Same School	-	-	-	<1% (18)	<1% (19)	-	-	-	11% (883)	9% (785)	-	-	-	67% (5,533)	64% (5,580)	-	-	-	22% (1,824)	27% (2,379)	-	-	-	100% (8,258)	100% (8,763)
Moved to a NewLocation	-	-	-	1% (10)	1% (9)	-	-	-	20% (197)	18% (134)	-	-	-	65% (656)	64% (476)	-	-	-	14% (145)	17% (129)	-	-	-	100% (1,008)	100% (748)
Total (N)	**-**	**-**	**-**	**28**	**28**	**-**	**-**	**-**	**1,080**	**919**	**-**	**-**	**-**	**6,189**	**6,056**	**-**	**-**	**-**	**1,969**	**2,508**	**-**	**-**	**-**	**9,266**	**9,511**

Source: TADS Feedback and Development Tool; TADS Student Performance Tool; HISD PeopleSoft Rosters: 2011–2012 as of 04–16–2012; 2012–2013 as of 04–10–2013; 2013–2014 as of 04–14–2014; 2014–2015 as of 05–15–2015; 2015–2016 as of 05–28–2016

†Accountability ratings not available for school year 2011–2012.

* Retention and Teacher Movement were not calculated in the TADS End of Year Reports for 2011–2012; 2012–2013; and 2013–2014

[1] 36 teachers excluded from data.

[2] 98 teachers at Community Services, HCC Life Skills, EL DAEP, or with no school identifying information in HR Roster. Not included in school levels.

[3] 108 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[4] 59 teachers without HR Roster identifying information.

[5] 78 teachers at Community Services, HCC Life Skills, EL DAEP, Beechnut Academy, or with no school identifying information in HR Roster. Not included in school levels.

[6] 57 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[7] 35 teachers without HR Roster identifying information.

Table E-2. 2011–2012 through 2015–2016 Instructional Practice (IP) Rating Distribution by Campus and Teacher Characteristics

	IP Level 1 (N)					IP Level 2 (N)					IP Level 3 (N)					IP Level 4 (N)					Totals (N)				
	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016	2011–2012	2012–2013	2013–2014	2014–2015	2015–2016
Overall Instructional Practice	2% (170)	2% (163)	2% (202)	1% (159)	1% (139)	11% (1,165)	16% (1,652)	14% (1,486)	13% (1,375)	12% (1,352)	70% (7,103)	65% (6,758)	65% (6,953)	64% (6,977)	63% (6,928)	17% (1,742)	17% (1,789)	20% (2,137)	22% (2,336)	24% (2,596)	100% (10,180)	100% (10,362)	100% (10,778)	100% (10,847)	100% (11,015)
School Level																									
Elementary	2% (86)	2% (99)	2% (101)	1% (89)	1% (77)	12% (679)	18% (1,041)	14% (862)	14% (857)	13% (796)	70% (3,993)	64% (3,629)	65% (3,857)	63% (3,768)	63% (3,823)	17% (941)	16% (936)	19% (1,136)	21% (1,233)	22% (1,360)	100% (5,699)	100% (5,705)	100% (5,956)	100% (5,947)	100% (6,056)
Middle	2% (37)	2% (40)	3% (54)	2% (30)	2% (33)	11% (193)	16% (276)	17% (301)	14% (253)	14% (247)	70% (1,185)	66% (1,142)	65% (1,175)	66% (1,202)	68% (1,230)	17% (283)	16% (279)	16% (286)	18% (325)	17% (312)	100% (1,698)	100% (1,737)	100% (1,816)	100% (1,810)	100% (1,822)
High	1% (33)	1% (23)	2% (43)	1% (28)	1% (21)	10% (236)	12% (268)	12% (271)	9% (206)	10% (241)	69% (1,562)	69% (1,590)	67% (1,560)	67% (1,556)	63% (1,488)	19% (422)	19% (430)	19% (452)	23% (522)	26% (619)	100% (2,253)	100% (2,311)	100% (2,326)	100% (2,312)	100% (2,369)
Combined	3% (14)	0% (0)	0% (0)	1% (10)	1% (8)	11% (57)	11% (67)	8% (51)	7% (48)	10% (67)	68% (363)	65% (397)	53% (357)	56% (383)	47% (324)	18% (96)	24% (144)	39% (258)	35% (239)	42% (291)	100% (530)	100% (608)	100% (666)	100% (680)	100% (690)
Total	**170**	**162**	**198**	**157**	**139**	**1,165**	**1,652**	**1,486**	**1,364**	**1,351**	**7,103**	**6,758**	**6,949**	**6,909**	**6,865**	**1,742**	**1,789**	**2,132**	**2,319**	**2,582**	**10,180**	**10,361**	**10,765**	**10,749** [1]	**10,937** [4]
Accountability Rating†																									
Improvement Required (IR)	-	3% (49)	4% (60)	3% (62)	2% (34)	-	24% (427)	25% (432)	22% (433)	22% (314)	-	64% (1,162)	62% (1,047)	67% (1,286)	64% (893)	-	9% (164)	9% (161)	8% (151)	12% (164)	-	100% (1,802)	100% (1,700)	100% (1,932)	100% (1,405)
Met Standard	-	1% (114)	2% (142)	1% (95)	1% (104)	-	14% (1,225)	12% (1,051)	11% (930)	11% (1,033)	-	65% (5,596)	65% (5,899)	64% (5,616)	63% (5,965)	-	19% (1,625)	22% (1,971)	25% (2,166)	25% (2,416)	-	100% (8,560)	100% (9,063)	100% (8,807)	100% (9,518)
Total	**-**	**163**	**202**	**157**	**138**	**-**	**1,652**	**1,483**	**1,363**	**1,347**	**-**	**6,758**	**6,946**	**6,902**	**6,858**	**-**	**1,789**	**2,132**	**2,317**	**2,580**	**-**	**10,362**	**10,763**	**10,739** [2]	**10,923** [5]
Index 1 Scores†																									
25 or Less	-	0% (0)	0% (0)	0% (0)	0% (0)	-	0% (0)	0% (0)	0% (0)	29% (11)	-	80% (44)	50% (6)	0% (0)	58% (22)	-	0% (0)	0% (0)	0% (0)	13% (5)	-	100% (44)	100% (6)	100% (0)	100% (38)
26 to 50	-	3% (14)	6% (39)	4% (52)	3% (28)	-	28% (114)	32% (223)	23% (294)	21% (209)	-	59% (238)	59% (416)	67% (872)	65% (649)	-	9% (38)	4% (25)	6% (74)	11% (112)	-	100% (404)	100% (703)	100% (1,292)	100% (998)
51 to 75	-	2% (106)	3% (136)	2% (97)	2% (92)	-	21% (1,128)	18% (955)	15% (883)	15% (922)	-	67% (3,605)	67% (3,617)	68% (4,072)	67% (4,057)	-	10% (527)	13% (698)	16% (940)	16% (999)	-	100% (5,366)	100% (5,406)	100% (5,992)	100% (6,070)
Greater than 75	-	1% (39)	1% (27)	<1% (8)	<1% (19)	-	9% (402)	6% (300)	5% (187)	5% (209)	-	63% (2,853)	63% (2,893)	57% (2,001)	56% (2,176)	-	27% (1,220)	30% (1,407)	37% (1,308)	38% (1,470)	-	100% (4,514)	100% (4,627)	100% (3,504)	100% (3,874)
Total	**-**	**159**	**202**	**157**	**139**	**-**	**1,644**	**1,478**	**1,364**	**1,351**	**-**	**6,740**	**6,932**	**6,945**	**6,904**	**-**	**1,785**	**2,130**	**2,322**	**2,586**	**-**	**10,328**	**10,742**	**10,788** [3]	**10,980** [6]
Core Foundation Teachers																									
Core	2% (120)	2% (124)	4% (275)	1% (119)	1% (112)	12% (826)	17% (1,207)	19% (1,403)	13% (1,114)	13% (1,101)	69% (4,749)	64% (4,462)	55% (4,013)	64% (5,351)	63% (5,402)	17% (1,175)	17% (1,186)	22% (1,564)	22% (1,826)	23% (2,030)	100% (6,870)	100% (6,979)	100% (7,255)	100% (8,410)	100% (8,645)
Non-Core	2% (50)	1% (39)	1% (51)	2% (38)	1% (27)	10% (339)	13% (445)	11% (396)	11% (250)	11% (250)	71% (2,354)	68% (2,296)	66% (2,318)	67% (1,594)	64% (1,502)	17% (567)	18% (603)	21% (755)	21% (496)	24% (556)	100% (3,310)	100% (3,383)	100% (3,520)	100% (2,378)	100% (2,335)
Total	**170**	**163**	**326**	**157**	**139**	**1,165**	**1,652**	**1,799**	**1,364**	**1,351**	**7,103**	**6,758**	**6,331**	**6,945**	**6,904**	**1,742**	**1,789**	**2,319**	**2,322**	**2,586**	**10,180**	**10,362**	**10,775**	**10,788** [3]	**10,980** [6]
Critical Shortage Teachers																									
Critical Shortage	2% (47)	2% (48)	2% (51)	1% (58)	1% (47)	13% (299)	16% (396)	14% (356)	12% (499)	12% (611)	71% (1,702)	67% (1,659)	67% (1,669)	66% (2,695)	62% (3,077)	14% (341)	15% (358)	17% (413)	20% (829)	25% (1,204)	100% (2,389)	100% (2,461)	100% (2,489)	100% (4,081)	100% (4,939)
Non-Critical Shortage	2% (123)	1% (115)	2% (151)	1% (99)	2% (92)	11% (866)	16% (1,256)	14% (1,129)	13% (865)	12% (740)	69% (5,401)	65% (5,099)	64% (5,282)	63% (4,250)	63% (3,827)	18% (1,401)	18% (1,431)	21% (1,724)	22% (1,493)	23% (1,382)	100% (7,791)	100% (7,901)	100% (8,286)	100% (6,707)	100% (6,041)
Total	**170**	**163**	**202**	**157**	**139**	**1,165**	**1,652**	**1,485**	**1,364**	**1,351**	**7,103**	**6,758**	**6,951**	**6,945**	**6,904**	**1,742**	**1,789**	**2,137**	**2,322**	**2,586**	**10,180**	**10,362**	**10,775**	**10,788** [3]	**10,980** [6]
Years of Experience																									
New Teacher	4% (29)	2% (39)	5% (91)	5% (58)	3% (41)	26% (198)	33% (515)	30% (596)	33% (412)	34% (393)	64% (482)	59% (918)	59% (1,166)	59% (732)	58% (677)	6% (42)	6% (92)	6% (117)	3% (43)	5% (62)	100% (751)	100% (1,564)	100% (1,970)	100% (1,245)	100% (1,173)
1-5 Years	1% (38)	1% (28)	1% (46)	1% (35)	1% (43)	10% (349)	13% (368)	11% (356)	11% (333)	12% (412)	73% (2,527)	68% (1,954)	67% (2,197)	69% (2,123)	67% (2,311)	16% (558)	18% (533)	20% (660)	19% (586)	19% (659)	100% (3,472)	100% (2,883)	100% (3,259)	100% (3,077)	100% (3,425)
6-10 Years	2% (37)	1% (27)	1% (22)	1% (24)	<1% (19)	10% (246)	12% (267)	9% (206)	10% (220)	8% (164)	70% (1,704)	69% (1,567)	64% (1,437)	63% (1,386)	63% (1,338)	18% (432)	18% (416)	25% (568)	26% (558)	28% (595)	100% (2,419)	100% (2,277)	100% (2,233)	100% (2,188)	100% (2,116)
11-20 Years	2% (37)	1% (33)	1% (20)	1% (29)	<1% (23)	11% (245)	14% (314)	10% (210)	10% (264)	9% (228)	67% (1,484)	64% (1,448)	65% (1,374)	63% (1,648)	61% (1,599)	20% (439)	21% (472)	24% (503)	26% (682)	29% (766)	100% (2,205)	100% (2,267)	100% (2,107)	100% (2,623)	100% (2,616)
Over 20 Years	2% (29)	2% (30)	2% (22)	<1% (11)	<1% (13)	10% (127)	13% (168)	10% (115)	8% (135)	9% (154)	68% (906)	63% (789)	65% (776)	64% (1,056)	59% (979)	20% (271)	21% (268)	24% (289)	27% (453)	31% (504)	100% (1,333)	100% (1,255)	100% (1,202)	100% (1,655)	100% (1,650)
Total	**170**	**157**	**201**	**157**	**139**	**1,165**	**1,632**	**1,483**	**1,364**	**1,351**	**7,103**	**6,676**	**6,950**	**6,945**	**6,904**	**1,742**	**1,781**	**2,137**	**2,322**	**2,586**	**10,180**	**10,246**	**10,771**	**10,788** [3]	**10,980** [6]

Source: TADS Feedback and Development Tool; TADS Student Performance Tool; HISD PeopleSoft Rosters: 2011–2012 as of 04–16–2012; 2012–2013 as of 04–10–2013; 2013–2014 as of 04–14–2014; 2014–2015 as of 05-15-2015; 2015–2016 as of 05-28-2016

* n < 5

†Accountability ratings not available for school year 2011–2012.

[1] 98 teachers at Community Services, HCC Life Skills, EL DAEP, or with no school identifying information in HR Roster were not included in school levels.

[2] 108 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[3] 59 teachers without HR Roster identifying information.

[4] 78 teachers at Community Services, HCC Life Skills, EL DAEP, Beechnut Academy, or with no school identifying information in HR Roster. Not included in school levels.

[5] 57 teachers at schools without accountability ratings or no school identifying information in HR Roster.

[6] 35 teachers without HR Roster identifying information.

Table E-3. 2012–2013 through 2015–2016 Student Performance (SP) Ratings by Campus and Teacher Characteristics

	SP Level 1 (N)				SP Level 2 (N)				SP Level 3 (N)				SP Level 4 (N)				Totals (N)			
	2012–2013	2013–2014	2014–2015	2015–2016	2012–2013	2013–2014	2014–2015	2015–2016	2012–2013	2013–2014	2014–2015	2015–2016	2012–2013	2013–2014	2014–2015	2015–2016	2012–2013	2013–2014	2014–2015	2015–2016
Overall Student Performance	19% (673)	15% (648)	12% (560)	4% (148)	19% (678)	16% (689)	25% (1,156)	8% (256)	39% (1,426)	39% (1,639)	28% (1,292)	20% (678)	24% (856)	30% (1,268)	35% (1,624)	67% (2,238)	100% (3,633)	100% (4,244)	100% (4,638)	100% (3,320)
School Level																				
Elementary	19% (480)	14% (434)	12% (385)	5% (94)	21% (528)	17% (521)	25% (789)	9% (176)	40% (1,006)	39% (1,157)	30% (940)	21% (439)	20% (517)	30% (891)	32% (1,008)	66% (1,358)	100% (2,531)	100% (3,003)	100% (3,122)	100% (2,067)
Middle	17% (144)	18% (166)	18% (128)	8% (41)	13% (108)	13% (118)	31% (224)	9% (45)	38% (327)	40% (364)	27% (197)	24% (123)	32% (272)	29% (268)	24% (172)	59% (305)	100% (851)	100% (916)	100% (721)	100% (514)
High	0% (0)	0% (0)	4% (18)	2% (9)	0% (0)	35% (9)	13% (62)	3% (17)	0% (0)	27% (7)	16% (78)	17% (97)	0% (0)	27% (7)	67% (315)	78% (434)	0% (0)	100% (23)	100% (473)	100% (557)
Combined	19% (47)	15% (45)	9% (29)	2% (4)	17% (41)	14% (41)	26% (81)	10% (17)	38% (93)	37% (110)	24% (77)	11% (19)	27% (67)	34% (101)	41% (129)	78% (138)	100% (248)	100% (297)	100% (316)	100% (178)
Total	**671**	**645**	**560**	**148**	**677**	**689**	**1,156**	**255**	**1,426**	**1,638**	**1,292**	**678**	**856**	**1,267**	**1,624**	**2,235**	**3,630**	**4,239**	**4,632** [1]	**3,316**
Accountability Rating																				
Improvement Required (IR)	31% (189)	33% (165)	29% (192)	12% (21)	28% (171)	24% (116)	37% (241)	9% (16)	32% (199)	33% (161)	21% (137)	25% (44)	9% (58)	10% (51)	13% (82)	55% (98)	100% (617)	100% (493)	100% (652)	100% (179)
Met Standard	16% (484)	13% (483)	9% (368)	4% (127)	17% (507)	15% (573)	23% (915)	8% (239)	41% (1,227)	39% (1,477)	29% (1,155)	20% (633)	26% (798)	32% (1,215)	39% (1,542)	68% (2,135)	100% (3,016)	100% (3,748)	100% (3,980)	100% (3,134)
Total	**673**	**648**	**560**	**148**	**678**	**689**	**1,156**	**255**	**1,426**	**1,638**	**1,292**	**677**	**856**	**1,266**	**1,624**	**2,233**	**3,633**	**4,241**	**4,632** [1]	**3,313** [2]
Index 1 Scores																				
25 or Less	43% (6)	100% (5)	0% (0)	17% (1)	36% (5)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	33% (2)	0% (0)	0% (0)	0% (0)	50% (3)	100% (11)	100% (5)	100% (0)	100% (6)
26 to 50	34% (39)	39% (83)	31% (135)	11% (18)	35% (40)	25% (53)	35% (153)	8% (14)	29% (33)	28% (59)	21% (93)	29% (50)	0% (0)	8% (18)	13% (56)	52% (88)	100% (112)	100% (213)	100% (437)	100% (170)
51 to 75	23% (430)	19% (408)	14% (357)	6% (90)	20% (381)	20% (413)	30% (748)	10% (143)	40% (754)	40% (854)	30% (736)	22% (327)	17% (315)	21% (438)	26% (635)	62% (921)	100% (1,880)	100% (2,113)	100% (2,476)	100% (1,481)
Greater than 75	12% (198)	8% (151)	4% (68)	2% (39)	16% (252)	12% (223)	15% (255)	6% (98)	39% (637)	38% (725)	27% (463)	18% (299)	33% (537)	42% (810)	54% (933)	74% (1,223)	100% (1,624)	100% (1,909)	100% (1,719)	100% (1,659)
Total	**673**	**647**	**560**	**148**	**678**	**689**	**1,156**	**255**	**1,424**	**1,638**	**1,292**	**678**	**852**	**1,266**	**1,624**	**2,235**	**3,627**	**4,240**	**4,632** [1]	**3,316**
Core Foundation Teachers																				
Core	19% (673)	15% (648)	13% (536)	5% (131)	19% (678)	16% (689)	27% (1,104)	9% (228)	39% (1,426)	39% (1,636)	30% (1,227)	23% (589)	24% (856)	30% (1,260)	31% (1,280)	63% (1,642)	100% (3,633)	100% (4,233)	100% (4,147)	100% (2,590)
Non-Core	0% (0)	0% (0)	5% (24)	2% (17)	0% (0)	0% (0)	11% (52)	4% (27)	0% (0)	0% (0)	14% (65)	12% (89)	0% (0)	78% (7)	71% (344)	82% (593)	0% (0)	100% (7)	100% (485)	100% (726)
Total	**673**	**648**	**560**	**148**	**678**	**689**	**1,156**	**255**	**1,426**	**1,636**	**1,292**	**678**	**856**	**1,267**	**1,624**	**2,235**	**3,633**	**4,240**	**4,632** [1]	**3,316**
Critical Shortage Teachers																				
Critical Shortage	19% (673)	21% (114)	9% (126)	4% (59)	19% (678)	13% (73)	23% (334)	8% (121)	39% (1,426)	35% (187)	31% (450)	23% (345)	24% (856)	31% (168)	37% (534)	64% (945)	100% (3,633)	100% (542)	100% (1,444)	100% (1,470)
Non-Critical Shortage	0% (0)	14% (534)	14% (434)	5% (89)	0% (0)	17% (616)	26% (822)	7% (134)	0% (0)	39% (1,451)	26% (842)	18% (333)	0% (0)	30% (1,099)	34% (1,090)	70% (1,290)	0% (0)	100% (3,700)	100% (3,188)	100% (1,846)
Total	**673**	**648**	**560**	**148**	**678**	**689**	**1,156**	**255**	**1,426**	**1,638**	**1,292**	**678**	**856**	**1,267**	**1,624**	**2,235**	**3,633**	**4,242**	**4,632** [1]	**3,316**
Years of Experience																				
New Teacher	28% (169)	24% (208)	23% (120)	7% (19)	23% (140)	22% (189)	36% (190)	13% (33)	38% (229)	37% (319)	25% (128)	24% (62)	11% (68)	17% (146)	16% (83)	56% (146)	100% (606)	100% (862)	100% (521)	100% (260)
1-5 Years	16% (181)	14% (197)	13% (187)	5% (54)	15% (169)	15% (208)	26% (362)	9% (88)	39% (433)	40% (552)	27% (380)	22% (222)	29% (325)	31% (423)	33% (461)	65% (665)	100% (1,108)	100% (1,380)	100% (1,390)	100% (1,029)
6-10 Years	18% (139)	13% (113)	13% (117)	4% (27)	18% (135)	14% (121)	24% (212)	7% (44)	39% (296)	36% (314)	27% (246)	19% (128)	24% (184)	37% (324)	36% (321)	70% (466)	100% (754)	100% (872)	100% (896)	100% (665)
11-20 Years	15% (120)	12% (92)	9% (97)	3% (29)	20% (161)	14% (105)	21% (245)	7% (61)	41% (327)	39% (298)	30% (343)	21% (177)	24% (196)	35% (264)	40% (456)	69% (585)	100% (804)	100% (759)	100% (1,141)	100% (852)
Over 20 Years	17% (60)	10% (38)	6% (39)	4% (19)	21% (72)	18% (66)	21% (147)	6% (29)	39% (135)	42% (155)	29% (195)	17% (89)	24% (83)	30% (110)	44% (303)	73% (373)	100% (350)	100% (369)	100% (684)	100% (510)
Total	**669**	**648**	**560**	**148**	**677**	**689**	**1,156**	**255**	**1,420**	**1,638**	**1,292**	**678**	**856**	**1,267**	**1,624**	**2,235**	**3,622**	**4,242**	**4,632** [1]	**3,316**

Source: TADS Feedback and Development Tool; TADS Student Performance Tool; HISD PeopleSoft Rosters: 2012–2013 as of 04-10-2013; 2013–2014 as of 04-14-2014; 2014–2015 as of 05-15-2015; 2015–2016 as of 05-28-2016

[1] Six teachers with SP missing from HR Roster file.

[2] Three teachers at schools without accountability ratings or no school identifying information in HR Roster.

Appendix F: Tests of Significance

Table F-1. One-Way Between Analysis of Variance of Teachers' Summative Rating by School Level Campuses, 2015–2016

Groups	N	Mean	Std. Deviation	F	p	η^2
Elementary	6,056	3.15	0.54			
Middle	1,822	3.06	0.52			
High	2,369	3.20	0.52			
Combined	690	3.34	0.59			
Total	10,937	3.15	0.54	F (3, 10,933) = 51.89	<0.001	0.01

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect size conventions for η^2 are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers located at Community Services, HCC Life Skills, EL DAEP, and Beechnut Academy (n=43) were not included. Teachers without HR identifying information (n=35) were not included.

Table F-2. Independent Sample *t* Test of Teachers Summative Ratings by Accountability Rating Campuses, 2015–2016

Groups	N	Mean	Std. Deviation	*t*	p	*g*
Improvement Required	1,405	2.93	0.54			
Met Standard	9,518	3.19	0.53			
Total	10,923	3.16	0.54	*t* (1, 10,921) = -17.12	<0.001	-0.49

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect sizes are calculated for independent sample t-tests using Hedge's g. Hedge's g provides a measure of effect size weighted to account for different sample sizes. Effect size conventions for Hedge's g are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers located at Not Rated (NR) schools (n=57) are not included in the table. Teachers without HR identifying information (n=35) are not included in the table.

Table F-3. One-Way Between Analysis of Variance of Teachers' Summative Rating by Index 1 Score Campuses, 2015–2016

Groups	N	Mean	Std. Deviation	F	p	η^2
25 or Less	38	2.91	0.54			
26 to 50	998	2.93	0.54			
51 to 75	6,070	3.06	0.52			
Greater than 75	3,874	3.36	0.50			
Total	10,980	3.15	0.54	F (3, 10,976) = 325.35	<0.001	0.08

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect size conventions for η^2 are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.

Table F-4. Independent Sample *t* Test of Teachers Summative Ratings by Core Foundation Status, 2015–2016

Groups	N	Mean	Std. Deviation	*t*	p	*g*
Core Foundation	8,645	3.15	0.54			
Non-Core Foundation	2,335	3.19	0.53			
Total	10,980	3.15	0.54	*t* (1, 10,978) = 3.23	.001	0.07

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect sizes are calculated for independent sample t-tests using Hedge's g. Hedge's g provides a measure of effect size weighted to account for different sample sizes. Effect size conventions for Hedge's g are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.

Table F-5. Independent Sample *t* Test of Teachers Summative Ratings by Critical Shortage Status, 2015–2016

Groups	N	Mean	Std. Deviation	*t*	p	*g*
Critical Shortage	4,939	3.17	0.53			
Non-Critical Shortage	6,041	3.14	0.54			
Total	10,980	3.15	0.54	*t* (1, 10,978) = -2.68	0.007	-0.06

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect sizes are calculated for independent sample t-tests using Hedge's g. Hedge's g provides a measure of effect size weighted to account for different sample sizes. Effect size conventions for Hedge's g are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.

Table F-6. One-Way Between Analysis of Variance of Teachers' Summative Rating by Teachers' Total Years of Experience, 2015–2016

Groups	N	Mean	Std. Deviation	F	p	η^2
New Teacher	1,173	2.79	0.53			
1 to 5 Years	3,425	3.12	0.52			
6 to 10 Years	2,116	3.23	0.52			
11 to 20 Years	2,616	3.24	0.52			
More than 20 Years	1,650	3.26	0.53			
Total	10,980	3.15	0.54	F (4, 10,975) = 189.52	<0.001	0.06

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect size conventions for η^2 are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.

Table F-7. Independent Sample *t* Test of Teachers Summative Ratings by Retention Status, 2015–2016

Groups	N	Mean	Std. Deviation	*t*	p	*g*
Retained	9,511	3.19	0.51			
Exited	1,469	2.90	0.66			
Total	10,980	3.15	0.54	*t* (1, 10,978) = -16.15	<0.001	-0.54

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect sizes are calculated for independent sample t-tests using Hedge's g. Hedge's g provides a measure of effect size weighted to account for different sample sizes. Effect size conventions for Hedge's g are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.

Table F-8. Independent Sample *t* Test of Teachers Summative Ratings by Teacher Movement Status, 2015–2016

Groups	N	Mean	Std. Deviation	*t*	p	*g*
Remained at the same school	8,763	3.21	0.50			
Moved to a new location	748	3.03	0.55			
Total	9,511	3.19	0.51	*t* (1, 9,509) = 8.77	<0.001	0.30

Source: Teacher Appraisal and Development F&D Tool, 2015–2016

Note: Effect sizes are calculated for independent sample t-tests using Hedge's g. Hedge's g provides a measure of effect size weighted to account for different sample sizes. Effect size conventions for Hedge's g are: 0.2 = small effect, 0.5 = medium effect, 0.8 = large effect. Teachers without HR identifying information (n=35) are not included in the table.