Department of Research and Evaluation
Austin Independent School District

AISD Pilot Teacher Appraisal System, 2011–2012

Multiple measures of teacher evaluation

Lindsay M. Lamb, Ph.D.
Lisa N. T. Schmitt, Ph.D.

January 2013

TABLE of CONTENTS.

EXECUTIVE SUMMARY. In Spring 2011, a working group of teachers, principals, central office administrators, and representatives from Education Austin worked together to replace Austin Independent School District's (AISD) Professional Development and Appraisal System (PDAS; Texas's teacher evaluation system) with a new appraisal system that was piloted at three REACH schools (Sunset Valley Elementary School, Webb Middle School, and Lanier High School) during the 2011–2012 school year. To explore multiple methods of teacher evaluation, teachers at these schools received ratings from student course feedback surveys (based on the revised classroom observation rubric used in the pilot teacher appraisal system) and participated in the peer observation program, in which peer observers used the revised classroom observation rubric to evaluate teachers. This report analyzes pilot teacher appraisal data with multiple measures of teacher effectiveness (e.g., 2011 PDAS data, peer observation data, student course feedback data, and student achievement data) along with feedback gathered from teachers participating in the pilot appraisal system in Spring 2012. On average, teachers scored 81 out of 100 possible points on the pilot appraisal, with scores ranging from 56 to 98. Teachers received the highest percentage of possible points for administrators' observations, and the lowest percentage of possible points for school value-added gains. Teachers' 2012 pilot appraisal scores were higher, on average, than their 2011 PDAS scores. Pilot appraisal scores for elementary and middle school teachers were higher compared with scores for high school teachers. Scores, however, did not vary drastically for core compared to non-core teachers, despite non-core area teachers' concerns with the pilot teacher appraisal system. Teachers who met both student learning objectives (SLOs) also had higher administrator walk-through observation scores and total appraisal scores than did teachers who did not meet their SLOs (all levels). Although these results are tentative at best due to the small sample size, the data suggest that peer, administrator, and student course feedback data are valuable tools that can be used with SLOs and value-added data to provide a holistic evaluation of teachers within AISD.

HOW WERE TEACHERS EVALUATED IN THE PILOT TEACHER APPRAISAL SYSTEM?

As part of the pilot teacher appraisal system, teachers were evaluated using six different measures across four domains considered to represent effective teaching practices (Figure 1).

Domain 1: Student growth. Student growth was assessed using either value-added computations at the teacher level, where available, or results from their individually established SLOs. Teachers also were evaluated based on the results of their team SLOs (or second individual SLO) and school-wide value-added growth computations. It should be noted that SLO computations were computed as the percentage of teachers' students who made their SLOs rather than whether teachers met or did not meet their SLO targets. Forty percent of each teacher's appraisal was based on student growth.

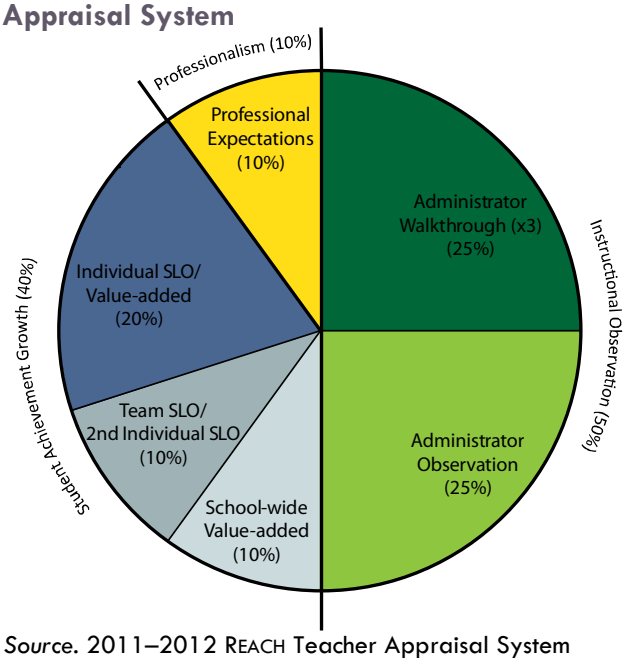
Domain 2: Instructional practice. Instructional practice was assessed using a classroom observation rubric that was redesigned for the pilot appraisal system (See Appendix A). Using the redesigned rubric, administrators conducted three walk-through observations and one formal classroom observation for each teacher.

Domain 3: Classroom climate. Classroom climate was assessed by administrators during the formal observations and walk-through observations using the revised observation rubric. Together, administrator observations constituted 50% of a teacher's total appraisal score.

Domain 4: Professional expectations. Professional expectations were assessed by administrators using the revised classroom observation rubric and evidence gathered from teachers throughout the academic year. Professional expectations accounted for 10% of the final appraisal score.

Figure 1 represents the percentage of points that each domain contributed to a teacher's total appraisal score. In the pilot teacher appraisal system, 60% of a teacher's total appraisal score was based on his or her administrators' ratings and 40% of his or her score was based on student performance. Components of the appraisal were converted so scores summed to 100 possible points. Scores less than 40 were considered *unsatisfactory*, scores between 40 and 59 were considered *developing*, scores between 60 and 79 were considered *effective*, and scores greater than 80 were considered *highly effective*. One hundred and one teachers had complete pilot teacher appraisal data (i.e., 23 elementary school teachers, 17 middle school teachers, and 61 high school teachers). Data were removed for one teacher with scores outside the range of possible appraisal scores.

Figure 1. Components of AISD's Pilot Teacher Appraisal System



Source. 2011–2012 REACH Teacher Appraisal System handbook

BOX 1. STUDENT GROWTH INDICATORS IN THE 2011–2012 PILOT TEACHER APPRAISAL SYSTEM. In addition to 39 instructional practice, classroom climate, and professional expectations competencies measured during administrator walk-through observations and formal classroom observations (Appendix A), the pilot teacher appraisal system includes multiple measures of student growth (i.e., SLOs and EVAAS). The table below describes each student growth measure used in the pilot teacher appraisal system, and the percentage of points allocated to each measure. Novice teachers and teachers placed on contractual difficultly were not evaluated using the new appraisal system.

Measure	Description	Grades and Subjects	Types of items/ assessment
Value-added growth (EVAAS) **(20% of appraisal)**	A district-rated measure of the extent to which students' average growth meets, exceeds, or falls short of average growth. Teacher comparison to state gain level will be used, and will be calculated by SAS EVAAS.	4–9 (Math and Reading) 5, 8, 9 (Science) 6,8,9 (Social studies)	STAAR exam
Individual student learning objective (SLO) **(20% of appraisal)**	Teachers set a target of student growth at the start of the school year and strive to achieve it by the end of the semester or school year. These targets are based on reviewing students' baseline skills and are set after approval from administrators.	All grades and subjects (except grades and subjects used in value-added computations)	Multiple choice/ performance based
Team student learning objective (SLO) or second individual SLO **(20% of appraisal)**	Teachers work together to set targets of student growth at the start of the school year and strive to meet by the end of the school year or semester. These targets are based on reviewing students' baseline skills by collaborating with team members and are set after approval from administrators. Teachers will each set and assess a team SLO or will utilize their second individual SLO.	All grades and subjects	Multiple choice/ performance
School-wide value -added **(10% of appraisal)**	A district-rated measure of the extent to which a schools' average growth meets, exceeds, or falls short of average growth. AISD has contracted with SAS EVAAS to compute this measure.	All schools	STAAR

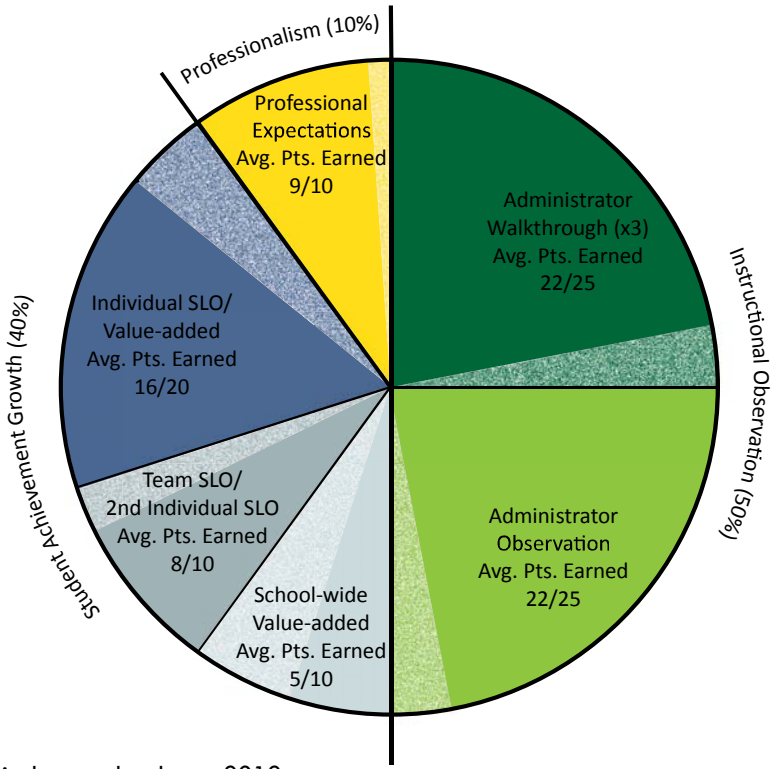
Source. 2011–2012 Teacher evaluation handbook
Note. This table was reproduced from the 2011–2012 Teacher evaluation handbook, which is no longer available online; please contact the authors for a copy.

"We went over [the new pilot appraisal system] in a staff development [meeting], but it was new and it was a lot of information in a 45-minute span…. To be honest, I don't really remember a lot." ~ REACH teacher, pilot teacher appraisal school

TEACHERS' PERFORMANCE ON THE PILOT TEACHER APPRAISAL SYSTEM

To understand the experience of teachers participating in the pilot appraisal system, 11 focus groups were conducted with a total of 50 teachers across participating schools. During the focus groups, teachers expressed some frustration and confusion with the pilot teacher appraisal system. Although participants received detailed information about the pilot teacher appraisal system at the beginning of the school year, they were unable to remember many details about the new system and were confused about who was eligible to participate in the pilot (e.g., novice teachers and teachers in contractual difficulty). Despite describing the pilot teacher appraisal system as "confusing," teachers on average received 81% of the total points possible on the pilot appraisal (Figure 2), thus scoring in the *highly effective* range. Teachers received the lowest percentage of possible points for the school-wide value-added component (5/10 possible points on average) and

Figure 2. Teachers' Average Scores for Each Component of AISD's Pilot Teacher Appraisal System. This figure represents the average number of points teachers received for each component of the appraisal.



Source. Teacher appraisal score database, 2012
Note. The solid area represents the average number of points teachers made in each component and the shaded area represents the average number of points teachers missed in each component.

received the greatest percentage of possible points for the administrator observations (22/25 possible points on both administrator observations and walk-throughs).

Performance by school level. Analyses were conducted to determine whether pilot appraisal scores varied for teachers across school level. It is important to remember that analyses across level only include schools included in the pilot appraisal system (i.e., one elementary school, one middle school, and one high school), and results may not generalize to the rest of the AISD population on the basis of school level.

Elementary and middle school teachers received higher total appraisal scores than did their peers at the high school level (Table 1), primarily due to differences in points received for value-added results. Middle and high school teachers received a greater number of points for the administrator walk-throughs than did elementary school teachers.

Analyses also were conducted to determine if teachers with individual value-added scores had different average appraisal scores than did teachers with individual SLO data. Results suggest no differences at the elementary or middle school level, but that high school teachers with individual SLO data received more points on component 2 of the appraisal than did those with individual value-added data (Table 1). Additionally, high school teachers with individual SLO data received significantly higher total appraisal scores (M = 80, with scores ranging from 57 to 95) than did high school teachers with value-added data (M = 69, with scores ranging from 56 to 80).

Table 1. Teachers' Average Scores in Each Component of AISD's Pilot Teacher Appraisal System, by Level (Elementary, Middle, and High School)

Component of the final appraisal	ES (n = 23)	MS (n = 17)	HS (n = 61)
1. Professional expectations (10 points)	9[a] (8–10)	9[b] (8–10)	8[a,b] (7–10)
2. Individual value-added/individual student learning objective (SLO) (20 points)	17 (5–20)	15 (9–19)	15 (4–20)
A. Individual value-added	16[a] (11–20)	15 (9–19)	11[a] (8–15)
B. Individual SLO	17 (5–20)	16 (9–19)	16* (4–20)
3. School value-added (10 points)	8[a] (8–8)	10[a] (10–10)	3[a] (3–3)
4. Team SLO/second individual SLO (10 points)	8 (4–10)	8 (0–10)	8 (1–10)
5. Administrator formal observation (25 points)	22 (19–26)	23 (20–25)	22 (13–26)
6. Administrator walk-throughs (25 points)	21[a,b] (18–24)	23[a] (19–26)	23[b] (18–26)
Total appraisal score (100 points)	83[a] (65–96)	88[b] (73–98)	78[a,b] (56–95)

Source. Teacher appraisal score database, 2012

Note. Means sharing the same superscript are significantly different from each other within the same component ($p < .05$).

The range of values for each component is presented in parentheses. Teachers with a perfect score on the administrator formal observation and walk-through could receive up to 26 points.

*indicates a significant difference between the number of points earned for individual value-added and individual SLO scores within the same grade level.

"[The pilot teacher appraisal system] feels like it is something that we have no control over." ~ REACH non-core area teacher, pilot teacher appraisal school

Performance by teacher type. During focus groups, non-core area teachers were concerned with the fairness of the pilot appraisal system for non-core area teachers (e.g., life skills, special education, physical education). For example, some non-core area teachers felt their SLOs, particularly their team SLOs, did not provide them enough opportunities to show or contribute to student growth.

BOX 2. TEACHER TYPE. Teachers were grouped as either core (e.g., those teaching reading, mathematics [math], science, and social studies; including bilingual and English as a second language teachers) or non-core (e.g., all remaining teachers).

As one non-core area teacher stated, "it is more difficult to quantify growth" in non-core area classes, particularly with certain student populations (e.g., fine arts classes, life skills students), potentially affecting these teachers' appraisal scores. Non-core area teachers described how their students often are pulled out of their classrooms to receive additional instruction, tutoring, or both, in other subjects, making it difficult for these students to meet their SLO targets and resulting in lower component 2 (individual value-added/SLO) scores.

Analyses were conducted to determine if the pilot appraisal system was biased against non-core area teachers. Examinations of pilot appraisal data for each teacher type (Table 2) suggest that scores were similar for core and non-core area teachers. However, non-core area teachers received slightly higher formal administrator observation scores than did core area teachers. Importantly, total

Table 2. Teachers' Average Scores in Each Component of AISD's Pilot Teacher Appraisal System, by Teacher Type

Component of the final appraisal	Core area teachers (n = 61)	Non-core area teachers (n = 40)
1. Professional expectations	9 (7–10)	9 (8–10)
2. Individual value-added/individual student learning objective (SLO)	15 (5–20)	16 (4–20)
3. School value-added	5 (3–10)	4 (3–10)
4. Team SLO/second individual SLO	8 (4–10)	8 (0–10)
5. Administrator formal observation	22 (13–26)	23* (20–26)
6. Administrator walk-throughs	22 (18–26)	22 (19–26)
Total appraisal score	80 (56–97)	83 (57–98)

Source. Teacher appraisal score database, 2012
Note. The range of values for each component is presented in parentheses next to the average score. Teachers with a perfect score on the administrator formal observation and walkthrough could receive up to 26 points.
* indicates a significant difference at $p < .05$

Table 3. Teachers' Average Scores in Each Component of AISD's Pilot Teacher Appraisal System, for Core and Non-core Area Teachers *With* and *Without* Value-added Data

Component of the final appraisal	Core teachers *without* value-added data (n = 39)	Core teachers *with* value-added data (n = 22)	Non-Core teachers (n = 40)
1. Professional expectations	9 (8–10)	8 (7–10)	9 (8–10)
2. Individual value-added/individual student learning objective (SLO)	16[a] (5–20)	13[a,b] (8–20)	16[b] (4–20)
3. School value-added	5 (3–10)	6 (3–10)	4 (3–10)
4. Team SLO/second individual SLO	8 (4–10)	7 (4–10)	8 (0–10)
5. Administrator formal observation	22 (17–26)	22 (13–26)	23 (20–26)
6. Administrator walk-throughs	22 (18–26)	22 (19–26)	22 (19–26)
Total appraisal score	81 (65–96)	77[a] (56–97)	83[a] (57–98)

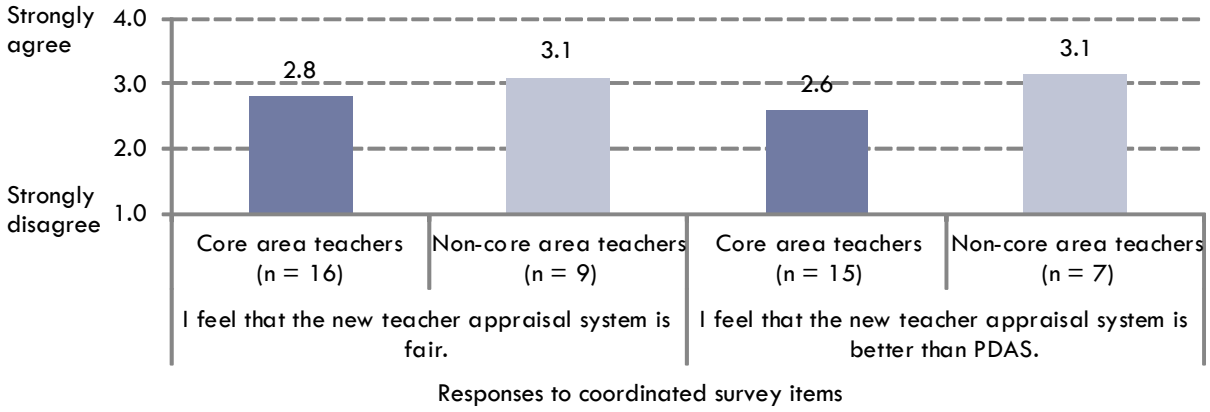
Source. Teacher appraisal score database, 2012
Note. The range of values for each component is presented in parentheses next to the average score. Teachers with a perfect score on the administrator formal observation and walkthrough could receive up to 26 points.

appraisal scores did not differ significantly between core and non-core area teachers.

When core teachers were divided into two groups (i.e., core area teachers *with* value-added data and core area teachers *without* value-added data) and compared with non-core area teachers, component 2 scores were lower for core teachers *with* value-added data than for both core area teachers *without* value-added data and non-core area teachers (Table 3). Additionally, core area teachers *with* value-added data received lower total appraisal scores than did non-core area teachers.

Examinations of survey data suggest that core and non-core area teachers had similar opinions regarding the fairness of the pilot appraisal system (Figure 3). That is, both core and non-core area teachers were more likely to agree than to disagree that the teacher

Figure 3. Teachers' Attitudes Towards the Pilot Teacher Appraisal System, Compared With the Professional Development Appraisal System (PDAS)



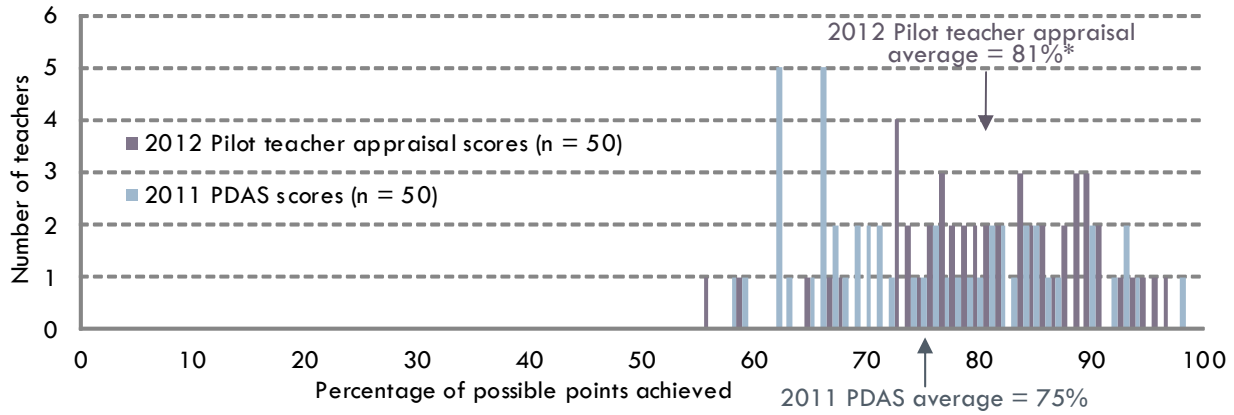
Source. 2012 Spring REACH ECS Survey
Note. Response options ranged from 4 (*strongly agree*) to 1 (*strongly disagree*), with scores greater than 3.0 considered desirable.

appraisal system was fair. Core-are teachers, however, were less likely to agree than non-core area teachers that the new teacher appraisal system was better than PDAS. Although non-core area teachers presented concerns during the focus groups, data suggest that the pilot appraisal system is not biased against non-core area teachers. In fact, the pilot appraisal system appears to be slightly biased against core area teachers *with* value-added data. As the new appraisal system expands to additional campuses, it will be important to continue monitoring these potential differences to ensure that the new appraisal system remains a fair measure of teacher performance for all teachers.

Pilot teachers' appraisal scores for 2012, compared with 2011 PDAS evaluations. Scores from the 2011–2012 pilot teacher appraisal system and 2010–2011PDAS scores were compared for those teachers who had data from both years (n = 50; Figure 4). On average, teachers earned a greater percentage of possible points on the pilot teacher appraisal (81%) than they earned on PDAS in 2011 (75%; results from t-tests are presented in Appendix B). High school teachers and non-core area teachers earned a greater percentage of points on the pilot teacher appraisal (78% and 86%, respectively) than they earned on PDAS in 2011 (70% and 73%, respectively). Finally, teachers with SLO data (as opposed to value-added data) included in their component 2 score received a greater percentage of points on the pilot teacher appraisal (83%) than they did on PDAS in 2011 (75%). It is important to note that novice teachers and teachers in contractual difficulty were evaluated with PDAS in 2012, not with the pilot teacher appraisal system. Thus, the sample of pilot appraisal teachers was limited to veteran teachers with previous satisfactory PDAS scores (i.e., those most likely to score well).

Although teachers were accustomed to administrators conducting classroom observations for PDAS, many participants did not report favorable experiences with administrator observations in the new appraisal. Specifically, focus group participants were concerned that they had not received all three walk-throughs from their

Figure 4. Distribution of 2010–2011 Professional Development Appraisal System (PDAS) and 2011–2012 Pilot Teacher Appraisal System Scores



Source. Teacher appraisal score database, 2012
Note. Scores were based on the percentage of possible points earned (100 on the pilot appraisal and 250 on PDAS).
* indicates percentage is significantly different at $p < .05$

administrator by the end of the school year or semester, that their administrator was not effective in reviewing their observation scores with them, and that the walk-throughs often were conducted too close together. Although the Office of Educator Quality encouraged administrators to conduct walk-throughs throughout the school year, some administrators were more effective than others in doing so. Regardless of these inconsistencies, teachers received favorable scores from administrators on both the formal observation and walk-through observation component of the new appraisal.

RELATIONSHIPS AMONG MULTIPLE MEASURES OF TEACHER EFFECTIVENESS

To explore multiple measures of teacher effectiveness, teachers' classroom instructional practices were evaluated by administrators (for the new appraisal), peer observers (for the strategic compensation program), and students (for a pilot student course feedback survey; see Box 3) at the three pilot appraisal schools. Peer observers and administrators used the same observation rubric to evaluate teachers (Appendix A), and students rated teachers using a survey with questions based on that rubric. Each observation measure evaluated the teachers' performance on the following domains: actively engaging students during instructional activities; checking for students' understanding and modifying instruction to address students' misconceptions; differentiating instruction for students' needs and using a variety of instructional strategies; developing problem-solving and critical-thinking skills for all students; setting rigorous academic expectations for students; providing relevant and successful feedback to students; setting and implementing classroom routines and procedures that support students' learning; establishing and maintaining standards for students' behavior; creating a safe and secure classroom environment that is organized and engages students; and establishing a climate that promotes fairness, respect, and diversity. Analyses were conducted to explore the potential relationships among the different observation measures and how they related to student achievement. Data also were examined to determine whether these measures were related to teachers' 2011 PDAS scores.

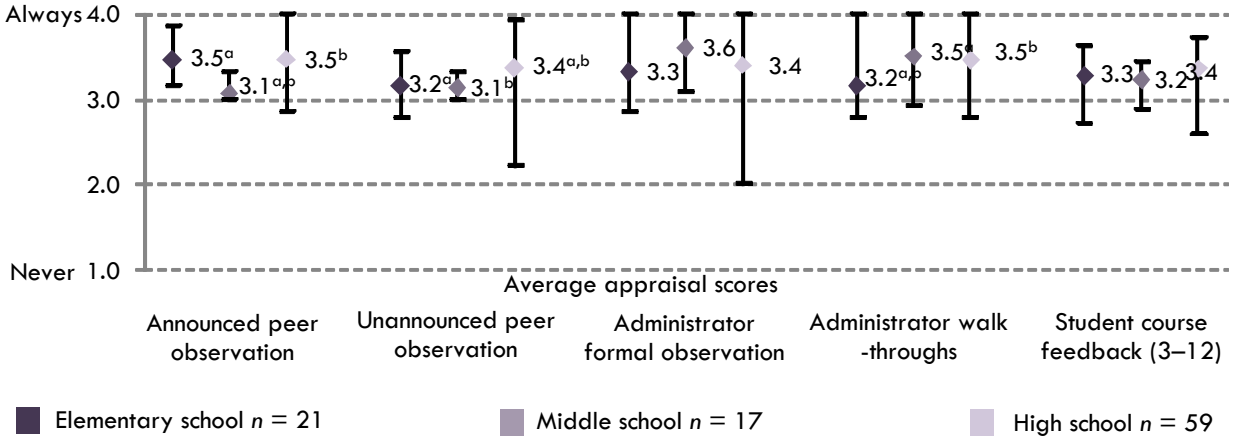
BOX 3. STUDENT COURSE FEEDBACK SURVEY PILOT. In addition to the peer and administrator observation metrics, in Spring 2011–2012, the Office of Educator Quality and the Department of Research & Evaluation piloted student course feedback surveys at the three pilot appraisal schools to determine if students' ratings of their teachers were similar to other observation metrics (e.g., peer observation scores and administrator observation scores) and if student feedback should be used in the new teacher appraisal system. Surveys were distributed to students in kindergarten through 12th grade. Previous analyses indicated that items were more reliable in grades 3 through 12 than in the earlier grades, and therefore are reported for those grades only in this report. Items on the student course feedback survey were based on the rubric developed for the peer observations and the administrator observations, but addressed specific indicators within domain and did not represent every aspect captured in each domain (for more detailed information, review the student course feedback **report**).

Teachers' average scores on the three observation measures. On average, teachers scored favorably on all three observation measures across school levels (Figure 5). Comparisons across school level suggest that high school teachers received significantly higher unannounced peer observation scores, on average, than did teachers at elementary and middle schools, and received higher administrator walk-through scores than did elementary school teachers. Elementary and high school teachers also received higher scores than did middle school teachers on the announced peer observation. Scores among high school teachers were more varied than were scores at the elementary and middle school level. However, teachers' observation scores were similar and favorable for core and non-core area teachers (Appendix C). Although scores varied somewhat by level, analyses were conducted across school level and teacher type to increase the number of teachers included in the analyses.

Relations among classroom observation measures. Analyses were conducted to determine the relationships among the various observation measures used to evaluate teachers participating in the pilot appraisal system (i.e., peer observation ratings, administrator observation ratings, 2012 total pilot teacher appraisal scores, student course feedback ratings, and 2011 PDAS scores).

Announced and unannounced peer observation scores, administrator observations and walkthrough scores, and student course feedback scores were all positively related (Table 4). Not surprisingly, multiple ratings by the same observer generally were correlated more strongly with each other than with those from other observers. That is, administrator walk-through observation scores and administrator formal observation scores were highly correlated with each other (r = .63, p < .01; across all levels), and the announced and unannounced peer observation scores were highly correlated with each other (r=.59, p < .01). However, moderate correlations

Figure 5. Average Scores and Range of Scores Across the Three Classroom Observation Measures for Pilot Appraisal Participants, by Level



Source. Peer observation database, administrator observation database, and student course evaluations, 2012
Note. Means that share the same superscript within observation measure are significantly different from each other (p <. 05), with bars representing the range of scores for each observation measure.
Scores ranged from 1 (*never*) to 4 (*always*) and were averaged across domain.
Minimum group sizes are reported for each level; more observations were included for some measures.

Table 4. Correlations Between 2012 Average Teacher Observation Metrics and 2011 Professional Development and Appraisal System (PDAS) ratings for Pilot Appraisal Participants

	1	2	3	4	5	6	7
1. Announced peer observation (*n* = 48)	—	.59**	.35**	.30**	.45**	-.00	.07
2. Unannounced peer observation (*n* = 49)		—	.35**	.45**	.37**	.05	.18
3. Administrator observation (*n* = 50)			—	.63**	.56**	.41**	n/a
4. Administrator walk-through (*n* = 50)				—	.43**	.16	n/a
5. Student course feedback grades 3–12 (*n* = 41)					—	-.03	.22
6. 2011 average PDAS scores (*n* = 41)						—	.35*
7. 2012 total pilot appraisal score (*n* = 50)							—

Source. Peer observation database, administrator observation database, and student course feedback, 2012; 2011 PDAS scores

Note. Administrator data are included in the total appraisal score; correlations between these variables are excluded.

Minimum cell sizes are reported.

Peer observation, administration observation, and student course feedback scores ranged from 1 (*never*) to 4 (*always*) and were averaged across domains.

* $p < .05$; **$p < .01$

were found between administrator observation and student course feedback scores (r=.56, $p < .01$), between administrator walk-through scores and the student course evaluation scores ($r = .43$, $p < .01$), and between administrator walk-through scores and unannounced peer observation scores ($r = .45$, $p < .01$). Relationships between observation measures appeared stronger at the high school level than at the middle and elementary school levels (Appendix D). With the exception of total pilot appraisal scores, correlations between the multiple observation measures and 2011 PDAS scores were small.

In 2012–2013, two unannounced peer observations (one in the fall and one in the spring) will become part of the new appraisal system to add an additional measure of teacher effectiveness. Using unannounced peer observation as the benchmark by which to gauge the other observation metrics, correlations were computed at the domain level to determine which measures achieved similar results (Table 5). These relationships offer evidence for moderate consistency across the observation measures at the domain level. Stronger relations may not be found, or even expected, between student and peer observation ratings due to the fact that peer observations were designed to evaluate teachers at a specific point in time whereas student course feedback scores reflected students' experiences with their teacher throughout the duration of their course or grade. Similarly, unlike peer observers' ratings, administrators' ratings might be influenced by their prior knowledge and experiences with teachers.

"One snapshot is not fair. It is good to have more observations." ~ REACH teacher, asked about including peer observations in the appraisal.

Table 5. Correlations Between Average Unannounced Peer observation scores and other classroom observation measures for pilot appraisal participants, by domain

Unannounced peer observation domain scores	Announced peer observation domain score (*n* = 92)	Administrator formal observation domain score (*n* = 58)	Administrator walkthrough domain score (*n* = 34)	Student course feedback 3–12 domain score (*n* = 74)
Actively engage students during instructional activities. (Domain 1)	.34**	.31*	-.09	.11
Checks for student understanding and modifies instruction to address student misconceptions. (Domain 2)	.30**	.38**	-.17	.27**
Differentiates instruction for student needs, utilizing a variety of instructional strategies. (Domain 3)	.22*	.17	.32	.04
Develops problem solving and critical thinking skills for all students. (Domain 4)	.30**	-.06	-.06	.15
Sets rigorous academic expectations for students. (Domain 5)	.29**	-.19	.03	.13
Provides relevant and useful feedback to students. (Domain 7)	.31**	.28*	.02	.29**
Sets and implements classroom routines and procedures that support student learning. (Domain 9)	.32**	.47**	.32	.20
Establishes and maintains standards for student behavior. (Domain 10)	.38**	.44**	.38*	.30**
Creates a safe and secure classroom environment that is organized and engages students. (Domain 11)	.39**	.52**	.16	.30**
Establishes a climate that promotes fairness, respect, and diversity. (Domain 12)	.31**	.18	-.03	.26*

Source. Peer observation database, administrator observation database, and student course evaluations, 2012
Note. Domains 6, 8, and 13 are excluded because they were considered less reliable.
Scores ranged from 1 (*never*) to 4 (*always*) and were averaged across domain.
Minimum cell sizes are reported.
* $p < .05$; ** $p < .01$

"Teachers would like to see something from the students' perspective [in their appraisals]." ~ REACH teacher, pilot teacher appraisal school

As the pilot appraisal expands, it will include both student course feedback scores and peer observation scores, along with administrator observation scores. In addition to the 50 teachers at the three participating pilot appraisal schools, focus groups were conducted with 155 teachers at 23 other schools that were implementing a new REACH peer observation program. During focus groups, teachers were asked how they felt about including student feedback scores and peer observation scores in their appraisal.

Most teachers expressed no concern about adding student course feedback scores to their appraisal. Indeed, some teachers were excited about the idea of including student feedback. However, one teacher said teachers were concerned that students "could be mad at you that day and give you a bad evaluation," and another said that students "don't have the maturity to deal with the [student course evaluation] and understand the consequences [of their ratings]." At the secondary level, some teachers were concerned that adding student feedback to teacher appraisals would discourage teachers from taking disciplinary actions because they feared students would retaliate with a low evaluation. Other concerns mentioned by elementary school teachers included believing some students "misunderstood the questions" and did not understand the differences between "always, often, and never." Yet other teachers were concerned that students would evaluate teachers based on popularity rather than on teacher quality, and as a result felt student feedback should be excluded from teacher appraisals. Despite these concerns, student feedback will be included as another component of the new teacher appraisal system in 2012–2013. Importantly, analyses in this and other reports (Schmitt, 2012) suggest that student feedback scores are reliable and

Table 6. Relationship Among Teachers' Average Scores on the Observation Metrics, Total Appraisal Score, and Student Learning Objectives (SLOs) for Pilot Appraisal Participants

	Did not meet SLOs	Met team SLO only	Met individual SLO only	Met two SLOs
Announced peer observation	3.36 (7)	3.35 (10)	3.54 (13)	3.38 (63)
Unannounced peer observation	3.16 (7)	3.24 (11)	3.22(13)	3.31 (67)
Administrator formal observation	3.26 (8)	3.19[a] (11)	3.44 (13)	3.47[a] (69)
Administrator walkthrough	3.14[a] (8)	3.38 (11)	3.31 (13)	3.46[a] (69)
Student course feedback (grades 3–12)	3.34 (6)	3.15 (10)	3.41 (9)	3.34 (50)
Total appraisal score	71%[a,b] (8)	73%[c] (11)	81%[b] (13)	83%[a,c] (69)

Source. Peer observation database, administrator observation database, and student course evaluations, 2012
Note. Means that share the same superscript are significantly different ($p < .05$) from each other within the same observation type (*n* counts are included in parentheses).
Scores ranged from 1 (*never*) to 4 (*always*) and were averaged across domain.

are related to other observation measures included in the pilot teacher appraisal.

Reactions to the possibility of including the peer observation ratings as part of the teacher appraisal system were mixed but the majority of teachers, though a small majority, preferred that the peer observations not be included in the appraisal system. Two primary reasons emerged in support of this preference. First, the teachers felt that the supportive and constructive nature of the peer observations would be compromised if the ratings were also used as evaluation. They felt the peer feedback is more valuable without the evaluative component. However, a notable number of teachers felt that the peer observations *should* be part of the appraisal system. Most of these teachers felt that that peer and administrator observations should be combined. They noted that the peer could be more objective, could add a different perspective, and that it would be fairer to have more than one observer rating their teaching skills.

Relationship between teacher observation measures and student outcomes. Because SLOs were included in the pilot appraisal as a measure of effectiveness, analyses were conducted to determine if teachers' scores on the various observation measures related to achieving their respective SLOs (Table 6). Teachers who met two of their SLOs had higher administrator walk-through scores than did teachers who did not meet either SLO. Similarly, teachers who met two of their SLOs had higher administrator formal observation scores than did teachers who only met their team SLO. Additionally, teachers who met two of their SLOs had higher total appraisal scores than did teachers who did not meet either of their SLOs. However, scores for announced and unannounced peer observations and for student course feedback did not differ for teachers who met SLOs and those who did not.

Finally, correlations were conducted between teachers' observation scores and value-added gains (Tables 7 and 8). Due to differences in value-added computations and data distributions across school level and subject area, analyses were computed separately by school level and subject area, where possible. In addition, data for teachers with peer observation data are included for REACH teachers with and without pilot appraisal data due to the data availability and the inclusion of peer observation as part of the revised pilot appraisal for 2012–2013. It is important to remember that data were still limited due to the number of teachers with value-added data (i.e., core-area teachers in grades 4 through

Table 7. Correlations Between Average Teacher Observation Scores and 2012 Subject Area Value-added Scores, by Level (Elementary, Middle, and High School)

	Math			Reading			Science			Social studies	
	ES	MS	HS	ES	MS	HS	ES	MS	HS	MS	HS
Announced peer observation average	.07 (95)	.01 (36)	.51 (26)	.17 (109)	-.47 (33)	.34 (24)	.24 (33)	—	.14 (23)	—	.40 (21)
Unannounced peer observation average	.09 (95)	-.06 (36)	.57 (24)	.11 (108)	-.27 (32)	.10 (23)	.24 (34)	—	.13 (22)	—	.22 (21)

Source. Peer observation database, 2012; 2012 EVAAS data
Note. Due to small cell sizes, *p* values were not considered meaningful and results should be interpreted with caution.
Cells are reported where $n >/= 10$ and — indicates the cell size is < 10 (*n* counts are included in parentheses).

Table 8. Correlations Between Teacher Average Observation Scores and 2012 Teacher Value-added Level Gains, by Subject Area for Pilot Appraisal Participants

	Math	Reading
Announced peer observations	-.06 (10)	.43 (11)
Unannounced peer observations	.00 (10)	.48 (12)
Administrator observation average	.20 (10)	.45 (11)
Administrator walk-throughs	-.14 (10)	.21 (11)
Student course feedback (4–12)	—	.19 (11)

Source. Peer observation database, administrator observation database, and student course evaluations, 2012; 2012 EVAAS data

10); therefore, results should be interpreted with caution.

Announced and unannounced peer observation scores were positively related to value-added gains in math, and announced peer observation scores were positively related to value-added gains in social studies for high school teachers. At the middle school level, however, announced peer observation scores were negatively related to gains in reading.

To compare the magnitude of relationships between peer observations, administrator observations, and student course evaluation data with value-added scores, correlations were computed for teachers at pilot appraisal schools. Due to sample size limitations, results were examined for reading and math only, and data were collapsed across levels (Table 8). Results suggest relationships between announced and unannounced peer observation scores and administrator observation scores with value-added level achieved in reading. Student course feedback also was moderately related to value-added level gains in math.

BOX 4. VALUE-ADDED LEVELS. To account for differences in the distribution of value-added scores by level, teachers were divided into five groups according to their scores (level 1 through level 5, with 5 indicating greater growth).

Due to the small number of teachers included in these analyses, however, results should be interpreted with caution.

CONCLUSIONS

Although some teachers expressed concerns about the pilot teacher appraisal system during focus groups, teachers scored favorably on the new appraisal system. Scores on the pilot appraisal were moderately related to student outcomes and other classroom observation measures, including peer observation and student course feedback scores. In 2012–2013, unannounced peer observation scores and student course feedback scores will be included in each pilot teacher's appraisal score. Importantly, student course feedback scores were positively related to administrator observation scores. Additionally, peer observation measures were positively related to measures of student outcomes in 2011–2012. Specifically, peer observation scores were positively related to value-added gains in math and social studies at the high school level. Components of the pilot appraisal were also related to SLOs. For example, teachers who met two of their SLOs had higher administrator walkthrough scores and total pilot appraisal scores than did teachers who did not meet either of their SLOs. Although these results are tentative at best due

to the small sample size, the data suggest that peer, administrator, and student course feedback data are valuable tools that can be used with SLOs and value-added data to provide a holistic evaluation of teachers within AISD.

REFERENCES

Lamb, L. M., & Schmitt, L. S. (2012). AISD *REACH program update, 2011–2012: The peer observation program* (DRE Publication No. 11.64). Austin, TX: Austin Independent School District.

Schmitt, L. S. (2012). *Student course feed back pilot, 2011–2012*. (DRE Publication No. 11.76RB). Austin, TX: Austin Independent School District.

APPENDIX

Appendix A. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Actively engages students during instructional activities	Students are rarely engaged during direct instruction as indicated by not completing instructional activities, unresponsive to questions, inaccurate following of teacher directions, and not asking appropriate questions Ineffectively manages time and materials Provides one way to engage lesson content or does not provide opportunities for student engagement Explains concepts and instructions in an incoherent and ineffective manner Students are disengaged during independent and group work Does not provide a balance between direct instruction and student-centered learning so that students are not able to apply their learning	Students are somewhat engaged during direct instruction as indicated by some completion of instructional activities, some responsiveness to questions, following of some of teacher directions, and asking of some questions Manages time and materials that is effective occasionally Provides two ways to engage lesson content Explains concepts and provides instructions but the delivery is not well organized or effective for learning Students display an effort to meet minimum expectations in independent and group work Provides more direct instruction than student-centered learning but students are able to apply their learning	Students are actively engaged during direct instruction as indicated by completion of instructional activities, responsiveness to questions, accurate following of teacher directions, and asking of appropriate questions Effectively manages time and materials Provides multiple ways to engage lesson content Clearly explains concepts and provides clear instructions Students display active engagement in independent and group work Provides a balance between direct instruction and student-centered learning so that students are able to apply their learning	Students are actively engaged during direct instruction as indicated by taking a leadership role in completing instructional activities and communicating lesson content to each other Students engage in learning at a consistent pace and know how to access all materials Students generate enrichment discussions or extensions in response to the lesson content Clearly explains concepts and provides clear instructions in a way that actively involves students in the learning process Students demonstrate enthusiasm in lesson activities in independent and group work Students take a leadership role during the lesson and require little guidance from the teacher when appropriate to maintain the pace and momentum

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Checks for student understanding and modifies instruction to address student misconceptions	Infrequently or never checks for understanding Does not diagnose misunderstandings and misconceptions or the checks are ineffective in accurately assessing student understanding Does not adjust lessons to ensure student understanding in response to assessments during the lesson or attempts to address the misunderstanding are ineffective Does not answer student questions or respond to their needs and interests during a lesson. Does not reteach.	Checks for understanding using a variety of methods some of the time Accurately diagnoses misunderstandings and misconceptions and responds with appropriate strategies some of the time Adjusts lessons to ensure student understanding in response to assessments during the lesson but interrupts the flow of the lesson Occasionally answers student questions or responds to their needs and interests during a lesson or reteaches content when necessary	Routinely checks for understanding using a variety of methods Accurately diagnoses misunderstandings and misconceptions and responds with appropriate strategies Adjusts lessons to ensure student understanding in response to assessments during the lesson and without interrupting the flow of the lesson Answers student questions or responds to their needs and interests during a lesson or reteaches content when necessary	Implements a system of checking for understanding that result in a substantive awareness of students' progress and needs in a lesson Anticipates student misunderstandings and misconceptions and preemptively addresses them Uses clear systems and routines for assessing student understanding during the lesson Answers student questions or responds to their needs and interests during a lesson or reteaches content when necessary

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Differentiates instruction for a student need utilizing a variety of instructional strategies	Does not adapt depth, pace, and delivery of what is taught to ensure all students have access to the lesson and are appropriately challenged Does not provide additional interventions, supports, enrichment, or variation of work in order to meet the needs and learning style of each student Does not provide multiple ways to engage lesson content Designs content, processes, lessons, and assessments that are applicable to only one group of students with different needs and interests	Adapts depth, pace, and delivery of what is taught to ensure all students have access to the lesson and are appropriately challenged some of the time Occasionally provides additional interventions, supports, enrichment, or variation of work in order to meet the needs and learning style of each student Provides two ways to engage lesson content Designs content, processes, lessons, and assessments that are applicable to some groups of students with different needs and interests	Adapts depth, pace, and delivery of what is taught to ensure all students have access to the lesson and are appropriately challenged Provides additional interventions, supports, enrichment, or variation of work in order to meet the needs and learning style of each student Provides multiple ways to engage lesson content Designs content, processes, lessons, and assessments that are applicable to sub-groups of students with different needs and interests	Students demonstrate clear systems that adapt the depth, pace, and delivery of what is taught based on their needs Students know their needs and actively seek learning experiences or tasks that meet their needs and learning styles Knows each student's level and needs and offers individualized instruction that is challenging to all students Works with students to design learning experiences and tasks to address their different needs and interests

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Develops problem-solving and critical thinking skills for all students	Creates limited situations or no opportunities to challenge students to think independently or critically about the content Does not develop and use various instructional strategies that challenge students and provide opportunities to engage in problem solving and critical thinking about the content Does not probe for higher-level thinking through questioning strategies Does not provide time in the lesson for students to analyze, and draw their own conclusions	Creates situations that challenge students to think independently or critically about the content some of the time Develops and uses one or two instructional strategies that challenge students and provide opportunities to engage in problem solving and critical thinking about the content Probes for higher-level thinking through questioning strategies occasionally Provides very little time in the lesson for students to analyze, and draw their own conclusions	Creates situations that challenge students to think independently or critically about the content Develops and uses various instructional strategies that challenge students and provide opportunities to engage in problem solving and critical thinking about the content Probes for higher-level thinking through questioning strategies Provides time in the lesson for students to analyze, and draw their own conclusions	Creates situations that challenge students to think independently or critically and students communicate their reasoning processes Embeds problem-solving and critical thinking skills into the lesson so that mastery requires students to utilize these skills Students exhibit higher level thinking through their questioning of each other Provides time in the lesson for students to design and implement inquiries and problem solving to analyze and draw their own conclusions
	Sets rigorous academic expectations for students	Does not communicate rigorous academic expectations to students, parents/ guardians, and the community Does not know and understand student levels or challenge students to stretch beyond their limits Does not reinforce that all students can learn Students are not engaged in rigorous work	Inconsistently communicates academic expectations to students, parents/ guardians, and the community Knows and understands some student levels and challenges some students to stretch beyond their limits Reinforces that some students can learn Students are engaged in some rigorous work	Communicates rigorous academic expectations to students, parents/ guardians, and the community Knows and understands some student levels and challenges students to stretch beyond their limit Reinforces that all students can learn Students are engaged in rigorous work	Students can articulate the rigorous academic expectations and teacher communicates expectations to parents/guardians and the community Students know their levels and are provided with opportunities to challenge themselves Students express their ability to learn and their expectation to meet their goals Students are engaged in rigorous work

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Collects, tracks, and uses student data to develop lesson plans and assessments	Does not use assessment data to determine student strengths and weaknesses and to develop some objectives to lesson plans Does not track student progress toward meeting objectives Does not analyze student progress to modify lesson plans and objectives Does not use assessments to measure student mastery of standards and objectives and does not provide multiple ways students can demonstrate mastery	Uses assessment data to determine student strengths and weaknesses and to develop some objectives and some lesson plans Tracks some students' progress toward meeting objectives Analyzes some students' progress to modify lesson plans and objectives Uses assessments inconsistently to measure student mastery of standards and objectives and provides one or two ways students can demonstrate mastery	Uses assessment data to determine student strengths and weaknesses and to develop objectives and some lesson plans Tracks students' progress toward meeting objectives Analyzes students' progress to modify lesson plans and objectives Routinely uses assessments to measure student mastery of standards and objectives and provides multiple ways students can demonstrate mastery	Embeds a wide array of ongoing assessments in lessons to determine student strengths and weaknesses and to develop objectives and lesson plans Students track their own progress toward meeting objectives and goals Analyzes student progress to modify lesson plans and objectives Routinely uses assessments to measure student mastery of standards and objectives that are interwoven into each lesson and provides multiple ways students can demonstrate mastery

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Provides relevant and useful feedback to students	Students do not engage in self-assessment Students do not understand assessment criteria and how they will be evaluated Provides incorrect feedback or no feedback during a lesson Does not provide feedback to students that explains why they have or have not met the standard	Some students engage in self-assessment Students understand some assessment criteria and how they will be evaluated some of the time Provides limited feedback during a lesson that affirms correctly understood content, clarifies misunderstood content, and extends student thinking. Provides feedback to students that explains why they have or have not met the standard some of the time.	Students engage in self-assessment Students understand assessment criteria and how they will be evaluated Provides feedback during a lesson that affirms correctly understood content, clarifies misunderstood content, and extends student thinking Provides high-quality, timely feedback to students that explains why they have or have not met the standard	Students self-assess on their understanding of lesson objectives and provide feedback to the teacher Students use assessment criteria to guide their learning Provides feedback during a lesson that affirms correctly understood content, clarifies misunderstood content, and allows students to synthesize concepts Provides high-quality, timely feedback to students that allows them to explain why they have or have not met the

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Instructional practice	Designs effective objective driven lessons and assessments that reflect the standards	Plans units and lessons that are frequently not aligned to state standards and district curriculum requirements Does not sequence lessons to ensure student mastery of standards and objectives Does not select, create, or adapt materials and resources to enrich learning Objectives of the lessons are not clear to students Does not develop rigorous lessons plans and activities that are developmentally appropriate and emphasize key concepts Lessons are not well organized and do not provide time for students to master objectives and standards	Plans units and lessons that are generally aligned to state standards and district curriculum requirements Sequences lessons to ensure student mastery of most standards and objectives Selects, creates, or adapts some materials and resources to encourage learning Objectives of the lessons are clear to some students Develops lessons plans and activities that are appropriate and emphasize some key concepts Lessons are somewhat organized and do not provide very little time for students to master objectives and standards	Plans rigorous units and lessons that are aligned to state standards and district curriculum requirements Sequences lessons to ensure student mastery of standards and objectives Selects, creates, or adapts some materials and resources to enrich learning Objectives of the lessons are clear to students Develops rigorous lessons plans and activities that are appropriate and emphasize key concepts Lessons are well organized and do not provides time for students to master objectives and standards	Plans rigorous units and lessons that are aligned to state standards and district curriculum requirements and have cross-curricular connections to key concepts Sequences lessons to help students synthesize and apply knowledge extending mastery of standards and objectives Selects, creates, or adapts materials and resources to extend student understanding Objectives of the lessons are clear to students and they can explain how the lesson connects to prior knowledge or learning Develops rigorous lessons plans and activities that demonstrate in depth student understanding of key concepts and relationships among various concepts and themes Lessons are well organized and provide time for students to master objectives and standards through continual engagement and self-assessment and reflection

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Classroom climate	Sets and implements classroom routines and procedures that support student learning	Does not design or implement consistent classroom routines and procedures that run smoothly Does not use instructional time so that students are engaged from the beginning of class to the end of class Fosters limited student independence or no independence through inefficient classroom routines and procedures	Designs and implements classroom routines and procedures but does not implement them consistently or teach them to students Uses instructional time so that some students are engaged from the beginning of class to the end of class Fosters some student independence through some shared responsibilities for classroom routines and procedures	Effectively designs and implements consistent classroom routines and procedures that run smoothly Effectively uses instructional time so that some students are engaged from the beginning of class to the end of class Fosters student independence through shared responsibilities for classroom routines and procedures	Effectively designs and implements consistent classroom routines and procedures that incorporate student responsibility and run smoothly Students assume responsibility for utilizing instructional time from the beginning of class to the end of class Students assume responsibility for routines and procedures and carry them out in an efficient manner with little or no direction from the teacher
	Establishes and maintains standards for student behavior	Does not clearly communicate high student behavioral expectations Does not reinforce appropriate behavior as needed Does not follow through on consequences Does not encourage or reinforce positive behavior Does not address off-task or inappropriate behavior efficiently Off-task or inappropriate behavior interferes with student learning	Communicates some student behavioral expectations Reinforces some appropriate behavior Inconsistently follows through on consequences Encourage or reinforce positive behavior inconsistently Address off-task or inappropriate behavior inconsistently Off-task or inappropriate behavior does not interfere with student learning some of the time	Clearly communicates high student behavioral expectations Reinforces appropriate behavior as needed Consistently follows through on consequences Encourages and reinforces positive behavior Address off-task or inappropriate behavior efficiently Off-task or inappropriate behavior does not interfere with student learning	Student demonstrate high behavioral expectations through their actions and require little redirection from the teacher Students hold each other accountable for appropriate behavior Consistently follows through on consequences Students encourage and reinforce positive behavior Addresses off-task or inappropriate behavior efficiently Off-task or inappropriate behavior does not interfere with student learning or does not occur

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Classroom climate	Creates a safe and secure classroom environment that is organized and engages students	Classroom is not a safe learning environment. Learning is accessible to some students. Class arrangement is not conducive to learning and does not change as needed for lessons Classroom environment does not display student work and exemplars Students do not have access to appropriate resources and technology Students are not invested in their work and do not value their academic success Students are not able to take risks and challenge themselves	Classroom is a safe learning environment. Learning is accessible to most students. Class arrangement is conducive to learning but does not change as needed for lessons Classroom environment displays some student work or exemplars Students have access to some resources and technology Students are invested in some of their work and sometimes show that they value their academic success Some students are able to take risks and challenge themselves	Classroom is a safe learning environment. Learning is accessible to all students. Class arrangement is conducive to learning and changes as needed for lessons Classroom environment displays student work and exemplars Students have access to resources and technology Students are invested in their work and value their academic success Students are able to take risks and challenge themselves	Classroom is a safe learning environment. Learning is accessible to all students. Class arrangement is a resource that is conducive to individual and group learning and students are able to contribute to the changing design of the environment Classroom environment displays student work and exemplars Students have access to resources and technology Students are invested in their work and value their academic success as shown through their ownership of classroom routines and procedures Students openly take risks and challenge themselves during class

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
	Establishes a climate that promotes fairness, respect, and diversity	Students do not actively listen or respond positively to each other and the teacher Teacher does not communicate or model expectations for respect of student differences Teacher does not have a positive rapport with students and does not ensure that all students contribute and their opinions' are valued Teacher does not celebrate student accomplishments	Students listen occasionally and respond to each other and the teacher Teacher communicates and models expectations for respect of student differences some of the time Teacher has a rapport with students and ensures that some students contribute and their opinions' are valued Teacher celebrates some student accomplishments	Students actively listen and respond positively to each other and the teacher Teacher communicates and models expectations for respect of student differences Teacher has a positive rapport with students and ensures that all students contribute and their opinions' are valued Teacher celebrates student accomplishments	Students take initiative socially and participate in creating a climate of respect Students demonstrate respect for student differences and encourage positive peer interactions Teacher develops a positive, caring rapport with students and ensures that all students contribute and their opinions' are valued Students celebrate each others accomplishments
Classroom climate	Provides responsive communication to parents throughout the year	Does not communicate with parents/guardians regarding performance, behavior, and school activities Does not promptly respond to parents/guardians within 1-2 school days Does not celebrate with parents/ guardians academic and social success Does not maintain a communication log Does not engage parents in students' academic success	Communicates infrequently with parents/guardians regarding performance, behavior, and school activities Responds to parents/guardians Celebrates with some parents/ guardians academic and social success Maintains a sparse communication log Engages some parents in students' academic success	Regularly communicates with parents/guardians regarding performance, behavior, and school activities Promptly responds to parents/ guardians within 1-2 school days Celebrates with parents/guardians academic and social success Maintains a communication log Engages parents in students' academic success	Regularly communicates with parents/guardians regarding performance, behavior, and school activities that communication results in changes to student behavior Promptly responds to parents/ guardians within 1-2 school days Regularly celebrates with parents/guardians academic and social success Maintains a thorough communication log Engages parents in students' academic success

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Professional expectations	Complies with district and school policies and procedures	Does not follow district and school policies Does not dress appropriately and professionally Does not comply with special education, 504, G/T, and ELL policies and procedures Does not arrive on time to school and to classes Does not prepare plans when absent or does not secure a substitute Does not implement school rules Teacher does not keep accurate records	Follows some district and school policies Dresses appropriately and professionally some of the time Complies with some special education, 504, G/T, and ELL policies and procedures Arrives on time to school and to classes most of the time Prepares minimal plans when absent and secures a substitute Implements school rules some of the time Teacher keeps some accurate records	Follows district and school policies Dresses appropriately and professionally Complies with special education, 504, G/T, and ELL policies and procedures Arrives on time to school and to classes Prepares plans when absent and secures a substitute Consistently implements school rules Teacher keeps accurate records	Follows district and school policies and invests effort to help them succeed Dresses appropriately and professionally Complies with special education, 504, G/T, and ELL policies and procedures and invests efforts to help them be successful students Arrives on time to school and to classes Prepares plans when absent and secures a substitute Consistently implements school rules and finds innovative ways to help them succeed Teacher keeps accurate records
	Fulfills professional responsibilities while modeling professional integrity	Does not participate in school and district policies Does not maintain a positive and productive relationship with colleagues, students, and parents and does not interact with all in a respectful manner	Participates in a few school and district activities Maintains productive relationships with colleagues, students, and parents with minor expectations	Participates actively in school and district activities Maintains a positive and productive relationship with colleagues, students, and parents and interacts with all in a respectful manner	Participates actively in school and district activities Maintains a positive and productive relationship with colleagues, students, and parents and promotes respect and professionalism

Appendix A, Continued. Administrator and Peer Observation Rubric for Walk-throughs and Formal Observations, 2011–2012

Domain	Competency	Level 1 description	Level 2 description	Level 3 description	Level 4 description
Professional expectations	Establishes professional goals, participates in professional development, and applies learning to practice	Does not identify professional development opportunities to enrich instructional practice Participates inconsistently or not at all in professional learning to improve student achievement Does not set professional goals or reflect on practice Does not welcome feedback from supervisors and colleagues in order to improve	Identifies professional development opportunities Participates and implements professional learning to improve student achievement with frequent reminders and monitoring Sets professional goals but does not reflect on practice Welcomes some feedback from supervisors and colleagues in order to improve	Identifies professional development opportunities to enrich instructional practice Participates and effectively implements professional learning to improve student achievement Sets professional goals and reflects on practice Welcomes feedback from supervisors and colleagues in order to improve	Identifies professional development opportunities to enrich instructional practice Leads professional learning that impacts student achievement or mentors others to effectively implement professional learning to improve student achievement Sets professional goals and reflects on practice Solicits feedback from supervisors and colleagues in order to improve
	Engages in meaningful collaboration to attain school goals and a positive campus climate	Does not collaborate with colleagues or share ideas and lessons to improve practice Does not support or contribute to campus goals and initiatives Does not actively participate in team/department meetings or school committees Does not collaborate with colleagues and administration to support struggling students	Collaborates with some colleagues or share ideas and shares a few ideas and lessons to improve practice Supports and contributes to some campus goals and initiatives Participates in some team/department meetings or school committees Collaborates with some colleagues and administration to support some struggling students	Collaborates with colleagues and shares ideas and lessons to improve practice Supports and contributes to campus goals and initiatives Actively participates in team/department meetings or school committees Collaborates with some colleagues and administration to support struggling students	Initiates or leads collaboration and the sharing of ideas and lessons to improve practice Supports and contributes to campus goals and initiatives Assumes leadership roles in team/ department meetings or school committees Initiates or leads collaboration with some colleagues and administration to support struggling students

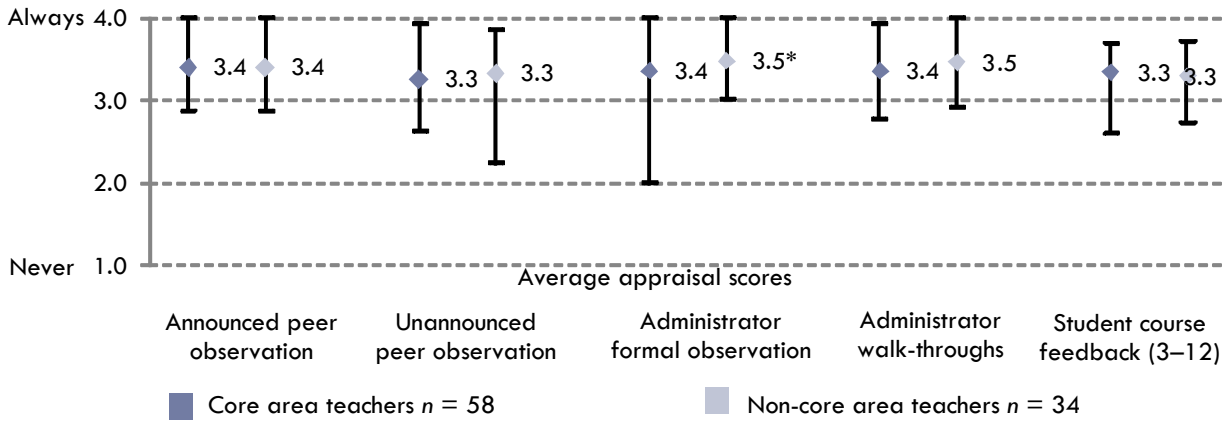
Source. 2011–2012 Teacher evaluation handbook observation rubric

Appendix B. Results For Paired *t*-tests Examining Differences in the Percentage of 2011 Professional Development and Appraisal System (PDAS) Points Earned and the Percentage

Comparison group	% of 2011 PDAS	% of 2012 teacher appraisal	*t*
All teachers (*n* = 50)	75%	81%	-3.73**
Elementary school teachers (*n* = 7)	83%	79%	1.46
Middle school teachers (*n* = 16)	81%	87%	-1.77
High school teachers (*n* = 27)	70%	78%	-4.58**
Core area teachers (*n* = 32)	77%	79%	-1.17
Non-core area teachers (*n* = 18)	73%	86%	-5.71**
Individual value-added data for component 2 of pilot teacher appraisal (*n* = 37)	76%	77%	-.16
Individual SLO data for component 2 of pilot teacher appraisal (*n* = 13)	75%	83%	-4.85**

Source. Peer observation database, administrator observation database, and student course evaluations, 2012
Note. The percentage of points earned for each group are presented and n counts included in parentheses.

Appendix C. Average Scores and Range of Scores Across the Three Observation Measures for Pilot Appraisal Participants, by Teacher Type



Source. Peer observation database, administrator observation database, and student course evaluations, 2012
Note. * means are significantly different from each other within the same observation measure.
Minimum cell sizes are reported.
Scores for each domain ranged from 1 (*never*) to 4 (*always*) and were averaged across domain.

Appendix D. Correlations Between 2012 Average Teacher Observation Metrics and 2011 Professional Development and Appraisal System (PDAS) Ratings for Pilot Appraisal Participants, by Level

Elementary school (3–5)	1	2	3	4	5	6	7
1. Announced peer observation (4)	—	.55†	.28	.27	-.06	—	.35
2. Unannounced peer observation (4)		—	.59*	.37	.55	—	.35
3. Administrator formal observation (4)			—	.63*	.52	—	n/a
4. Administrator walkthroughs (4)				—	.39	—	n/a
5. Student course evaluation (3–5) (4)					—	—	—
6. 2011 average PDAS scores (3)						—	-.19
7. 2012 total pilot appraisal score (3)							—
Middle school	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1. Announced peer observation (15)	—	.77**	.29	-.09	.14	-.01	.46
2. Unannounced peer observation (15)		—	.33	-.08	.44	.06	.40
3. Administrator formal observation (15)			—	.25	.39	.48	n/a
4. Administrator walkthroughs (15)				—	.48	-.10	n/a
5. Student course evaluation (15)					—	-.15	.63*
6. 2011 average PDAS scores (14)						—	.05
7. 2012 total pilot appraisal score (15)							—
High school	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1. Announced peer observation (25)	—	.59**	.57**	.51**	.52**	.67**	.37**
2. Unannounced peer observation (26)		—	.36**	.50**	.31*	.44*	.44**
3. Administrator formal observation (27)			—	.67**	.65**	.54**	n/a
4. Administrator walkthroughs (27)				—	.46**	.67**	n/a
5. Student course evaluation (24)					—	.29	.34*
6. 2011 average PDAS scores (24)						—	.35
7. 2012 total pilot appraisal score (27)							—

Source. Peer observation database, administrator observation database, and student course evaluations, 2012; 2011 PDAS scores.
Note. Correlations are not included between administrator scores and total pilot appraisal scores because administrator scores are included in total appraisal score calculations.
Cells are reported where $n >/= 10$ and — indicates the cell size is < 10. Minimum cell sizes are presented.
Scores for each domain ranged from 1 (*never*) to 4 (*always*).