

Results in Brief: Study of Emerging Teacher Evaluation Systems

November 2016

Comprehensive teacher evaluation systems are a core element of current state and local strategies to support improved teaching and learning in elementary and secondary education in the United States. The Study of Emerging Teacher Evaluation Systems provides descriptive information on the design and early implementation of teacher evaluation systems in eight school districts located in Austin, TX; Chattanooga (Hamilton County), TN; Colorado Springs (Harrison School District 2), CO; Tampa (Hillsborough County), FL; Pittsburgh, PA; Plattsburgh, NY; Leonardtown (St. Mary's County), MD; and Washington, DC. It is intended to provide lessons that inform the work of states and school districts in designing and implementing their own teacher evaluation and support activities, and to inform future research on the effects of teacher evaluation systems on teacher professional practice and student performance.

Study Questions

1. What key priorities and measures informed the design of the new teacher evaluation systems?
2. What steps did the districts take prior to full implementation to test the system and prepare teachers and staff to implement it?
3. How did the districts structure and conduct the classroom observation component of their teacher evaluation systems?
4. How did the districts analyze student performance and other data to evaluate teacher performance?
5. How did the districts use, or plan to use, teacher evaluation results to make personnel decisions? To what extent were professional development and career advancement decisions tied to evaluation results?
6. What administrative structures did districts use to support their new teacher evaluation systems?
7. What are the perceived early effects of the teacher evaluation systems on the professional practices of teachers, principals and district administrators?

Study Design and Limitations

This descriptive study relied on interviews and focus groups with district administrators, principals, teachers, and community members in eight case study districts. These districts were selected because their teacher evaluation systems relied on multiple measures of teacher performance, including classroom observations and measures of student performance; distinguished among teachers at various levels of proficiency and effectiveness; included a formative component that provided timely feedback to teachers; and used, or planned to use, evaluation results to inform decisions about teacher professional development, retention, tenure, and other personnel decisions.

Key study limitations include: (1) reliance on data collected from a purposively selected group of districts, thus limiting the range of potential successes and challenges districts might experience in designing and implementing complex teacher evaluation systems; (2) responses that may not represent the full range of views among district and school staff; and (3) a point-in-time review (i.e., from late 2012 to early 2013) of evaluation systems that will likely evolve.

Highlights

- Teachers and central office staff generally agreed that the foremost goal of the teacher evaluation system was to improve instruction.
- Teacher and principal input during the design and/or pilot test phase strongly influenced decisions regarding system modifications in six districts, according to district administrators.
- Classroom observations varied in frequency, duration, and degree of formality in all eight districts. In addition, principals reported challenges in finding time to conduct teacher observations.
- Six districts used multiple approaches for measuring teacher impact on student performance, including individual and/or school-level value-added models.
- Districts used teacher evaluation results for a range of purposes, including targeted professional development and support, career ladders and performance pay, and in some instances, redeployment or release of teachers identified as ineffective.
- The majority of districts created relatively simple, streamlined structures to administer their teacher evaluation systems.
- Teachers reported that they believed that the classroom observations and feedback helped them become better teachers.

Designing an Evaluation System

Many choices and tasks influence the design of a comprehensive teacher evaluation system that relies on multiple measures of performance to assess teacher effectiveness. Determining district purposes for developing the new teacher evaluation system was often the first step in the design process.

Charlotte Danielson's *Framework for Teaching* was the exclusive basis for the design of observation rubrics in half of the districts included in the study.

All eight districts used a range of assessments to measure teachers' influence on their students' performance. In addition to the state assessment, all districts used locally developed and/or standardized assessments (e.g., DIBELS, PSAT, AP, and IB).

To account for the many types of assessments and data sources used to measure teacher impact on student performance, seven districts assigned separate weights to each type of student assessment included in their evaluation system.

Early Implementation

The early implementation phase — pilot testing the system, introducing it to teachers, and training classroom observers — varied across the eight districts, yet may have been a critical juncture in determining the successful transition to a new, comprehensive teacher evaluation system.

Three of the eight districts did not pilot test their teacher evaluation systems before implementing them. Administrators in these districts explained that holding teachers accountable as early as possible was important for achieving real, measurable change in teaching and learning.

Each of the eight districts used one or more strategies to introduce teachers to the new evaluation system, including distributing written materials, creating online resources, or using school-based teams of evaluation experts.

Conducting Classroom Observations

Districts' approaches to conducting classroom observations varied with respect to the types of observers used, and the frequency, duration, and degree of formality of the observations.

The frequency of classroom observations was wide-ranging across the eight districts, from two annual observations for experienced teachers in one district to 18 observations for new teachers in another.

Three districts used peer observers, in addition to principals, to conduct classroom observations. Administrators in these districts explained that peer observers were intended to offer an unbiased, objective opinion of a teacher's performance.

Using Student Performance and Other Data

Districts used varied and changing strategies in their evaluation systems to measure teacher impact on student performance. Seven districts used student learning objectives among their methods for measuring teacher impact on student performance.

In four districts, some teachers who participated in focus groups or individual interviews criticized the use of student performance data to evaluate their performance, expressing concerns that the tests do not fully capture students' actual growth or are not aligned with the school curriculum.

Using Teacher Evaluation Results

Seven districts had defined, or begun to define, the results they would apply to teacher evaluation scores, including compensation, professional development, and other personnel decisions.

All eight districts used similar rating systems for classifying a teacher's overall performance.

Six districts used or planned to use teacher evaluation scores to redeploy or release low-performing teachers but narrowly defined the circumstances under which this could happen.

Three districts linked — or planned to link — teacher evaluation scores to a career ladder, which sometimes included performance pay.

Administering Evaluation Systems

Despite the complexity of the teacher evaluation systems, six of the eight districts created modest administrative structures with, on average, five staff administering their teacher evaluation systems.

Four districts worked with outside contractors to create data management systems for their teacher evaluation data, and two districts worked with contractors to train their classroom observers.

Hiring peer observers was among the most expensive features of the teacher evaluation systems in three districts that took this approach.

Perceived System Effects

Respondents at all levels reported perceiving positive effects of their new teacher evaluation systems, notably that observations and feedback had helped teachers improve their professional practice, even in the partially implementing districts.

Principals in six districts reported that the teacher evaluation system had caused them to want to be better instructional leaders.