# The Albuquerque Public Schools School Improvement Grant (SIG) Teacher Evaluation and Compensation Pilot

Sade Bonilla and Richard Bowman

Data Fellows

Albuquerque Public Schools

## Introduction

The federal School Improvement Grant (SIG) required districts to identify factors that assisted rigorous, transparent, and equitable evaluation systems. In order to satisfy this requirement, at SIG recipient sites, Albuquerque Public Schools (APS) conducted a Teacher Evaluation and Compensation pilot that utilized multiple observations by administrators, data on student growth, student learning goals, and student perception surveys. SIG grant funds were utilized to support performance-based compensation for participating staff members. Pilot participation did not supplant participation in the regular district evaluation process.

The major factor that drove the timeline for implementing the APS Teacher Evaluation and Compensation pilot was a requirement written into the SIG grant; however, the state of New Mexico's approved NCLB Waiver will require districts across the state to implement a multiple measures evaluation system starting with the 2013-14 school year. The chief academic officer (CAO) believed it was important to conduct an evaluation pilot in order to inform stakeholders and learn about various evaluation metrics before being required to implement them.

### Strategic Data Project Fellows' Roles

The Strategic Data Project (SDP) data fellows, Richard Bowman and Sade Bonilla, utilized existing research and policy developments in the field of teacher evaluation to develop a pilot that fit the local context. The data fellows subsequently facilitated the pilot process and worked collaboratively with the CAO, Linda Sink, and Albuquerque Teachers Federation (ATF) President, Ellen Bernstein, to guide pilot administration and policy communications. The data fellows administered student surveys, provided training for teachers and principals, developed the observation and student learning goal rubrics, designed and distributed teacher-level reports, served as liaisons between participants and the district, and conducted all data analysis as well as

the research reporting associated with the pilot.

## Local Context

Albuquerque Public Schools serves approximately 90,000 students K-12 at 13 high, 27 middle, 89 elementary, 10 alternative and 21 locally authorized charter schools. APS is one of the area's largest employers with approximately 12,000 employees. The pilot was conducted at four schools—three high schools and one middle school, that were identified as low-performing priority schools by the New Mexico Public Education Department (PED). The four schools serve a majority Hispanic population (81.5%) and have significantly higher levels of Free and Reduced Price Lunch (FRPL) (83.2%) and English Language Learners (ELLs) (22.6%) than the district averages, which are 65.2, 65.8, and 12.4 percent respectively.

## Project Vision

Data collection and analysis was a central focus of the pilot program. The data analysis and subsequent outcomes have served as a basis for determining policy recommendations for the district and state uses of the 'multiple measures' in the pilot. The data fellows saw an opportunity to utilize their knowledge of education policy, data analysis, and leadership skills to carry out a project that would assist the district and state of New Mexico understand several nationally proposed measures for evaluating teacher effectiveness. Improving teaching is a major focus of the CAO's Academic Plan. The pilot gave the CAO the ability to create and utilize a new teacher observation rubric, an event that supported an APS goal of improving teaching and learning.

## Getting Started

The data fellows were asked by the CAO to attend a presentation by a vendor to determine whether the vendors' product could satisfy the districts needs for teacher evaluation.

The data fellows determined that the vendor did not have any experience with teacher evaluation and was simply proposing to hire an outside contractor to analyze the districts original test scores. As a result, the data fellows recommended that the district utilize existing resources to design a teacher evaluation pilot and utilize the federal SIG grant funds to compensate teachers. The data fellows felt confident they could design and administer a system that would better suit the district's needs. After proposing a teacher evaluation pilot plan internally to the CAO, associate superintendent of secondary schools and assistant superintendent of human resources, the CAO and formed a steering committee with the ATF president, to negotiate the details of the pilot.

After negotiation, the members of the steering committee presented the finalized pilot components and timeline to the staff at each school site to solicit volunteers. The CAO and ATF president addressed the need for a new evaluation system, the importance of conducting a pilot to learn and gather feedback, and the political environment in the state and around the nation on implementing new teacher evaluation systems. The data fellows provided an overview of each pilot component and answered follow up questions at the school site presentations, as well as via email. Teachers and support staff at the school sites signed up for the pilot through their site administrator or the data fellows.

**Project Implementation**

The data fellows carried out most aspects of the pilot, relying on the CAO and ATF president for guidance on major decisions and contact with legislators, New Mexico Public Education Department (PED) officials, and media.

**Figure 1: Project Timeline for APS**

| **Project Timeline** | |
|---|---|
| **March, 2011** | • Data fellows present a proposal for a teacher evaluation pilot at an APS middle school. |
| **August 4, 2011** | • SDP Fellows attend a presentation by a vendor for a service to satisfy the districts teacher evaluation needs.<br>• SDP Fellows learn that the teacher evaluation pilot has expanded to include three high schools in addition to the middle school. |
| **August 8, 2011** | • Data fellows present an updated proposal for the teacher evaluation and compensation pilot to the CAO, associate superintendent of secondary schools, principal support specialists for secondary schools, and the director of research, deployment, and accountability. |
| **August 18, 2011** | • Data fellows meet privately with the CAO and cement her support for the pilot pending Board approval. |
| **August 26, 2011** | • Data fellows present the pilot proposal to the Board of Education and receive approval. |
| **September, 2011** | • The CAO includes the union president on the pilot steering committee.<br>• The steering committee discusses and negotiates the pilot's components and timeline, finalizing the pilot elements September 26th.<br>• Data fellows present to the principals of the four schools on September 30th. |
| **October, 2011** | • Data fellows present to school staff October 11, 12, 14, and 18.<br>• Data fellows conduct Q&A follow-up visits at each site on October 14, 18, 19, and 20.<br>• Participants sign up by the deadline of October 21. |

| | |
|---|---|
| **November, 2011** | • Data fellows conduct a pilot administrator orientation November 3 to review the new observation rubric tool and organize logistics for the teacher orientation.<br>• Participating teachers attend a voluntary half-day pilot orientation November 8-10 to review pilot components, observation frameworks, and the student learning goal rubric. |
| **December, 2011** | • Student surveys are administered at the end of the fall semester.<br>• Administrators submit teacher observation records for the first teacher observation. |
| **January, 2012** | • Classroom teachers receive reports detailing fall semester student survey results and interim value-added results.<br>• Participants submit Student Learning Goals via an online form developed by data fellows. |
| **February, 2012** | • Participants receive a personalized report detailing their pilot unit accumulation. |
| **March, 2012** | • Classroom teachers in tested subject receive interim value-added results from the winter administration of the district benchmark assessment. |
| **April, 2012** | • Student surveys are administered at the end of the spring semester. |
| **May, 2012** | • Individualized reports detailing student survey results and individual and school-wide value-added measures (VAM) estimates are distributed.<br>• Participants submit student learning goal results online.<br>• Participants submit self-ratings on the student survey and observation tool, and general pilot feedback.<br>• Principals submit the second and third teacher observation ratings. |
| **Summer, 2012** | • Participants receive final reports of individual and school-wide value added results and unit accumulation.<br>• Data fellows submit stipend sheets and participants receive compensation according to the units received for the various pilot components.<br>• Data fellows analyze results and produce reports for district leadership. |

**Key Decision Points**

The initial pilot proposal consisted of four measures of teacher performance: 1) a student perception survey, 2) multiple teacher observations, 3) individual student growth, and 4) school-wide student growth. The ATF union president was adamant that student learning objectives be added as an additional measure, due to her belief that student learning objectives were a valid way to measure student growth. The steering committee decided to include teacher-created goals despite the knowledge that due to the limited staffing and time, there would be little support, training or oversight of the creation of these goals. This addition resulted in two additional measures—individually and collaboratively-created goals—where participants would self-report results.

Much of the negotiation surrounding the pilot centered on whether all teachers in the school would participate or only a subset who volunteered. While the data fellows advocated for mandatory participation in order to model the most realistic system, the contentious nature of teacher evaluation and the negotiating interests of the union president resulted in the decision that participation would be voluntary.

**Key Wins**

There were two key short-term wins: first, an acceptance of the initial proposal presentation (for one school); second, gaining the trust of the union president and bringing her on-board with the design. The first win came by being prepared, organized and confidently showing that the overall project was feasible. The second win came from the willingness to negotiate some of the details of the project, and the ability to highlight the political climate to an advantage -- emphasizing the importance of having ‘skin in the game’ in order to have more political clout when commenting on teacher evaluation policy at the state level.

**Key Resistance Points**

Despite the fact that the pilot received union support, many teachers were skeptical of the pilot's motives. At one of the school sites a few teachers were openly hostile toward the data fellows and made inappropriate remarks. Teachers and principals at this school and one other school reported that teachers were experiencing "bullying" with regards to their decision to volunteer for the pilot. The data fellows visited each school site to provide a chance for teachers to ask follow-up questions about the pilot. Although this alleviated concerns at some school sites, the two schools with reports of "bullying" experienced significantly lower levels of staff volunteerism than the other two sites.

**Key Resources**

The data fellows relied on research in the education field on VAM and observation rubrics to inform the development of the pilot components. As part of the Strategic Data Project, the data fellows were able to gain access to the student perception survey utilized in The Measures of Effective Teaching Project. The student perception survey administration guide was a modified version of a survey administration guide utilized by Denver Public Schools, another SDP agency. Dan Goldhaber, the faculty advisor for APS, was consulted throughout the pilot to discuss potential research papers as well as policy implications and messaging strategies.

In order to carry out the pilot, the data fellows utilized Google Docs to collect information from teachers and principals including the observation rubric ratings, feedback surveys, and verification forms. The data fellows relied extensively on STATA to perform analysis and Excel and VBA Macros to create individualized reports for pilot participants.

**Key Skills Required**

1. The art of negotiation: Resulting in the addition of student learning goals, the weights of the multiple measures in a teacher's evaluation, and the buy-in of the union representatives.

2. Presentation skills: Effectively explaining the pilot components to a non-technical audience. Answering educators' questions about value added and survey analysis to alleviate their concerns that components were unfair.

3. Report making: Developing easy to understand reports for a non-technical audience on VAM and student perception survey results. Using technology (Excel and VBA Macros) to efficiently create and distribute individualized reports.

4. Change management: Displaying empathy by actively listening, responding, and validating individuals' fears and concerns.

5. Political savvy: Framing the pilot through multiple lenses which allowed various stakeholders to see the shared value in the project.

## Outcomes

The initial goal was to design a teacher evaluation plan for one middle school. This goal expanded into four schools, and evolved into the implementation and administration of a larger teacher evaluation system. The goal was ultimately accomplished. The data fellows are currently documenting the pilot implementation process and authoring research papers for the larger education policy community.

**Further Development and Potential Impact**

The overall project has been cited multiple times in the television, radio, and print news media and by district staff as an example of how the district is pushing reform forward. Parts of the pilot project have been incorporated into the operations of the agency - the district has signed a MOU with the union to use the observation protocol and associated tool for all teachers. The value-added model has been used on an as-needed basis to provide information to principals about their staff. It is reasonable to expect that the lessons learned in the project will be incorporated into the district's forthcoming individual Race To The Top (RTTT) grant application. Finally, the state Public Education Department (PED) has taken notice, and asked the data fellows and district to present the results to the New Mexico Teacher Evaluation Advisory Council which is providing guidance to the state about how to implement the new teacher evaluation system called for in their NCLB waiver application.

A significant body of research on teacher effectiveness has demonstrated that there are large differences in teaching effectiveness across teachers. Recent research (Taylor and Tyler, 2012) has documented that teachers who undergo an evaluation cycle with multiple observations and feedback improve after their evaluation. This evidence suggests that superior teacher evaluation systems have the ability to improve teaching quality. The wide variation of teacher effectiveness in APS and school districts across the nation suggests that successful efforts to improve teaching quality will have an impact on student achievement.

Teacher evaluation reform is at the forefront of national policy discussions. The federal government has encouraged states to re-invent their teacher evaluation systems through RTTT grant fund requirements and the NCLB Waiver process. This pilot was a learning process and will add to the research and documentation of best practices associated with evaluation systems

that utilize multiple measures, as well as contribute to the district's upcoming RTTT grant application.

## Lessons Learned

### Strengths

Developing and administering a new teacher evaluation without having a history in the organization requires paying close attention to relationships and feelings. Gaining buy-in from stakeholders is a key element for successful implementation; providing all people the chance to be heard while being respectful of their viewpoints will help garner that buy-in. Execution is important; thinking ahead and knowing the limits will help prevent over-promising. Have an answer for questions that will be asked – even if it is to say that the question is unanswerable or that there is no answer. Recognize and show appreciation to the people who must change their viewpoint or actions to move your project along. Maintain early, regular, and focused contact with stakeholders; the union president and the Chief Academic Officer's regular meetings about supporting quality teaching provided a supportive setting for the discussions about this pilot.

### Challenges

Start much earlier. Although a draft pilot plan was developed in early March of 2011, the data fellows were not informed that the plan would be utilized until after the school year had started in August. Although data fellows worked as quickly as feasible, due to the late notice, negotiations occurred throughout September and the pilot did not begin until October, two months into the school year, and therefore, principals missed one of the four originally planned observations.

The data fellows utilized existing resources to carry out all aspects of the pilot. The

district owned scanning software technology that allowed for paper-based administration of the student surveys. Over 13,000 student surveys were administered, all of which were printed and subsequently scanned in-house. The student survey had to compete with other district printing needs and resulted in student survey production and processing times that were somewhat greater than initially projected by the fellows.

## Recommendations for the District

The major recommendation for APS is to consider the ways to systematically use portions of the pilot (surveys, observations, VAM) to improve teaching and learning in the district. The implementation of the Common Core State Standards (CCSS) can serve as a catalyst for examining effective teaching practice. The district may consider using VAM estimates to develop lists of teachers eligible to be promoted into crucial out-of-classroom positions, or deliver professional development to others. APS currently has access to a library of high-quality videos of teaching that can be used to create online training modules for administrators and teachers. Administrators who participated in the pilot received little training on the observation rubric that was developed for the pilot. The next step to improving inter-rater reliability is to develop a training process for certifying observers.

Scaling up a multiple measure evaluation system will require higher levels of administrative personnel and an increase in technological capacity. The district may consider administering student surveys in an online format to ensure faster processing. Additionally, the district may consider developing an online form and storage capacity to collect the results of teacher observations to allow for simplified reporting and analysis.