



STRATEGIC **DATA** PROJECT

## SDP FELLOWSHIP CAPSTONE REPORT

# Colorado State Model Evaluation System for Teachers: Findings and Lessons Learned

**Britt Wilkenfeld, Colorado Department Of Education**

*SDP Cohort 3 Fellow*

### **SDP Fellowship Capstone Reports**

SDP Fellows compose capstone reports to reflect the work that they led in their education agencies during the two-year program. The reports demonstrate both the impact fellows make and the role of SDP in supporting their growth as data strategists. Additionally, they provide recommendations to their host agency and will serve as guides to other agencies, future fellows, and researchers seeking to do similar work. *The views or opinions expressed in this report are those of the authors and do not necessarily reflect the views or position of SDP or the Center for Education Policy Research at Harvard University.*

## **Introduction**

In 2010, Senate Bill 10-191 was passed changing the way teachers are supported and evaluated in the state of Colorado with the ultimate goal of ensuring college and career readiness for all students. To support school districts in implementing the new evaluation requirements, the Colorado Department of Education (CDE) is developing a model system as an option for districts to use for teacher evaluations. The Colorado State Model Evaluation System for teachers was piloted in 26 school districts of varying size and location during the 2012-13 school year. Quantitative and qualitative data from the pilot districts will inform changes to the state model system. The CDE is determined to have a research-based evaluation system that can be provided as a resource to be used by any district in the state. In my role as a Data Strategist at the CDE, I am responsible for collecting and analyzing all data relevant to the state model system, including evaluation ratings, baseline and feedback survey data, and other measures of classroom success (e.g., student growth scores, student perception surveys, school accountability measures, etc). In this capstone report and the accompanying Teacher Pilot Report I will describe the implementation process for this project and present preliminary findings. These analyses and findings are the first step in providing evidence for the reliability and validity of the Colorado State Model Evaluation System. As of September 2013, 161 of the 178 districts in the state of Colorado are using the SMES as their local evaluation system.

## **Project Vision**

Improving the effectiveness of all educators is one of the CDE's strategic goals, with the ultimate goal of improving students' outcomes. This meant that there was already agency-wide commitment to this work, and no-one in the CDE needed to be further convinced of the importance of the educator effectiveness work and the development of the Colorado State Model Evaluation System. However, some people did need some convincing of the importance of collecting and analyzing data, particularly collecting data. Districts in Colorado are strained on resources and every single data collection (new and old) places a burden on them. So it was important to explain to district representatives and state representatives what

data would be collected, how and when the data would be collected (giving them notice months ahead of time), and why the data would be collected. Following up on the “why,” I created an official report of the teacher pilot findings in order to provide a summary of initial findings and to show that the educator effectiveness team was indeed using the data that districts had worked so hard to submit to us (see the attached *2012-13 Teacher Pilot Report*). I have presented findings from this report to key stakeholders, including district and state leadership, policy leaders, and funders.

Although reporting pilot findings is important, the most important use of the data is informing changes and improvements to the state model system. Because 2012-2013 was the first year that the 26 districts used the state model system for teachers, there were obviously issues with training and implementation fidelity. So it is important to refrain from jumping to conclusions based on the data, but rather to flag findings to be explored further either through deeper quantitative analyses or qualitative analyses.

### **Project Implementation**

Collecting data from districts requires preparation and notification. I created and sent district representatives a 2012-2013 Pilot Data Collection Timeline in August 2012. Throughout the school year I had to be increasingly amenable to deadline extensions to accommodate for local priorities that districts were handling with in their local setting. This was obviously frustrating and it affected my own timelines, but it was important to give them extra time to complete different components of the process. I believe that the CDE was able to collect more data as a result of the flexible deadlines and districts were very appreciative of the department’s understanding and flexibility.

The other aspect that required planning was the mode of data collection. For some of the data collection (e.g., feedback surveys), I owned all aspects of the process so I merely needed to plan ahead (e.g., enter survey questions into an online survey response collector and collect email addresses from districts). For other aspects of the data collection (e.g., evaluation ratings), I had to work with the vendor who housed the CDE rubric. This partnership required constant contact and iterations with the vendor regarding what data

needed to be collected and how the data would be reported back to me. I reviewed many versions of spreadsheets to make sure I was given all of the variables I needed to conduct my own analyses.

## **Outcomes**

My main goal since starting this project has been to analyze the pilot data in order to inform improvements to the Colorado State Model Evaluation System (e.g., identify bias in the rubric, remove redundant elements, etc.). What I quickly realized is that the findings are too preliminary to inform substantial policy changes this early in the research; there are simply too many inconsistencies in training and implementation fidelity to attribute too much to the ratings.

I had to explain to the CDE and district leadership why the data were too preliminary to jump to conclusions because many of them were anxious to have conclusive findings. Instead, I've encouraged my team is using the findings to flag issues to investigate further, identify districts to target for increased support, and inform additional tools and resources that are needed in the field. As the research continues and I have more confidence in the data it will then be more appropriate to recommend any necessary changes in policy and practice.

Another goal has been to fully utilize the data that districts submit to the CDE and to report back to them in a useful and comprehensible way. I have been able to report out in the aggregate, but I would like to do more in terms of reporting at the school and district level. I am currently learning Tableau so that I can more easily produce reports for the educators and district leaders who took the time to participate in our pilot and submit data through focus groups, surveys, and evaluation ratings.

Originally I was not particularly interested in reporting the data out to larger audiences, but that has become one of my primary roles. In the fall of 2013 I presented pilot findings to the Colorado State Board of Education, State Council for Educator Effectiveness (a group of education leaders appointed by the Governor to monitor the implementation of SB-191), Colorado State Senator Mike Johnston, the Bill and Melinda Gates Foundation, and the U.S. Department of Education. It has been a good opportunity to expose others to the work

the educator effectiveness team at CDE has done and how thoughtful we are being in developing the state model system.

### **Lessons Learned**

It is important to discuss potential analyses and findings with like-minded data individuals in the agency. These conversations help ensure that the analyses are appropriate and comprehensive. One of the first groups that I presented unofficial findings to consisted of another data fellow, an agency fellow, and an analyst with expertise in data display. This group asked questions about the analyses I ran and pushed on elements of the data display to make sure I was accurately representing the findings.

It is also very valuable to discuss findings with content experts, meaning people in the agency with experience in the subject being analyzed (e.g., former teachers or principals). In my agency, I initially presented to four of my team members: the person who developed the rubric on which the ratings were based, two former teachers who trained district personnel on the state model system, and a leader on the team who is very interested in data. Only half of the graphs that I showed this group were included in the final report and presentation, as the other half were not easily interpreted by non-technical people. If the graphs or tables do not communicate the message then they are not worth keeping. This initial run-through of the findings was very valuable for helping me understand how the data and analyses were being understood or not by various audiences.

I next presented to my entire Educator Effectiveness team at the department. This experience helped me anticipate questions that might surface through my presentation and also helped me hone in on what people thought was important. Only after these three aforementioned presentations/discussions did I produce the final report on the teacher pilot findings; meaning the findings, presentation of data, and interpretations had already been vetted by colleagues with varying levels of content knowledge and technical expertise. The process led to a comprehensive but understandable report and accompanying presentation, both of which have been well- received by a range of audiences. I will follow the same process when analyzing data from the Colorado State Model Evaluation System for principals.