Stated Briefly

Redesigning teacher evaluations: Lessons from a pilot implementation

Julie Riordan
Karen Shakman
Quincy Chang
Education Development Center

Natalie Lacireno-Paquet
Candice Bocala
WestEd

In collaboration with the Northeast Educator Effectiveness Research Alliance

Based on a study (see box below) that examined the implementation of new teacher evaluation systems in New Hampshire, this brief provides information about five factors related to implementation: time and resource capacity of evaluators, training provided to evaluators, introduction and development of student learning objectives, stakeholder support from evaluators and teachers, and teachers' perceptions of professional climate in their schools.

Why this study

Federal policies and programs such as the Race to the Top grant program of 2009 and flexibility waivers of 2011 have required educator evaluation systems across the country to be redesigned. In 2009 only 14 states required annual teacher evaluations, but by 2012, 43 states required annual evaluations of all new

This brief presents selected findings from Riordan, J., Lacireno-Paquet, N., Shakman, K., Bocala, C., & Chang, Q., (2015), *Redesigning teacher evaluation: Lessons from a pilot implementation* (REL 2015–030), Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. That report is available at http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2015030.pdf.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL NORTHEAST & ISLANDS
Regional Educational Laboratory
At Education Development Center, Inc.

teachers (National Council on Teacher Quality, 2012). Research on the design and implementation of new teacher evaluation systems remains in the early stages. Although states are developing systems in different ways—for example, by providing general design guidelines to districts or mandating the implementation of specific models—some considerations are common to the design and implementation of all these systems.

Regional Educational Laboratory (REL) Northeast & Islands, in collaboration with the New Hampshire Department of Education and the Northeast Educator Effectiveness Research Alliance, studied the implementation of New Hampshire's new educator evaluation framework in 15 of the state's School Improvement Grant (SIG) schools located across eight districts during the 2012/13 school year.[1] This brief presents selected findings from this study about the factors related to implementation in these sites.

REL Northeast & Islands and the New Hampshire Department of Education collected and analyzed three sources of data for this study: district-developed evaluation plans and instruments; survey data from evaluators[2] and teachers; and interview data from district administrators, principals, and teachers. See box 1 for more information about sources of data and analytic approach.

Factors related to implementation

The analysis of teacher and evaluator survey data, as well as interviews with teachers, principals, and district administrators, suggested five factors that were especially salient to the implementation of New Hampshire's teacher evaluation framework in eight districts with a total of 15 SIG schools:

- Time and resource capacity of evaluators.
- Training provided to evaluators.
- Introduction and development of student learning objectives (SLOs).
- Stakeholder support from evaluators and teachers.
- Teachers' perceptions of professional climate in their schools.

Box 1. Data and methodology

Administrative documents and data. The study team collected district plans and instruments used for teacher evaluation. District plans contained information on the features of each district's teacher evaluation system and processes.

Surveys. The New Hampshire Department of Education developed two online surveys: an evaluator survey for principals and other evaluators and a teacher survey for teachers in the schools with School Improvement Grants (SIGs). The surveys gathered data on evaluator and teacher perceptions and experiences with the evaluation systems. The overall response rate was 88 percent for the evaluator survey (n = 35) and 61 percent for the teacher survey (n = 277).

Interviews. The study team conducted semistructured interviews with a small sample of district administrators (n = 5), principals (n = 8), and teachers (n = 6) from the SIG schools to supplement survey findings.

Analytic approach. The study team collected data for all eight districts with SIG schools in New Hampshire and analyzed evaluation plans and other documents from these districts. The study team then compared the documented features against the reported use of the features from teacher surveys to create an index of implementation fidelity for each district. To analyze factors related to implementation, the study team and the New Hampshire Department of Education designed interviews and surveys drawing on the specific questions from the department as well as the research literature on implementation, including time and personnel, planning and training, stakeholder support, and leadership. The study team analyzed survey responses on evaluator and teacher perceptions about the new evaluation system and used interview data to supplement the survey analysis.

Time and resource capacity of evaluators

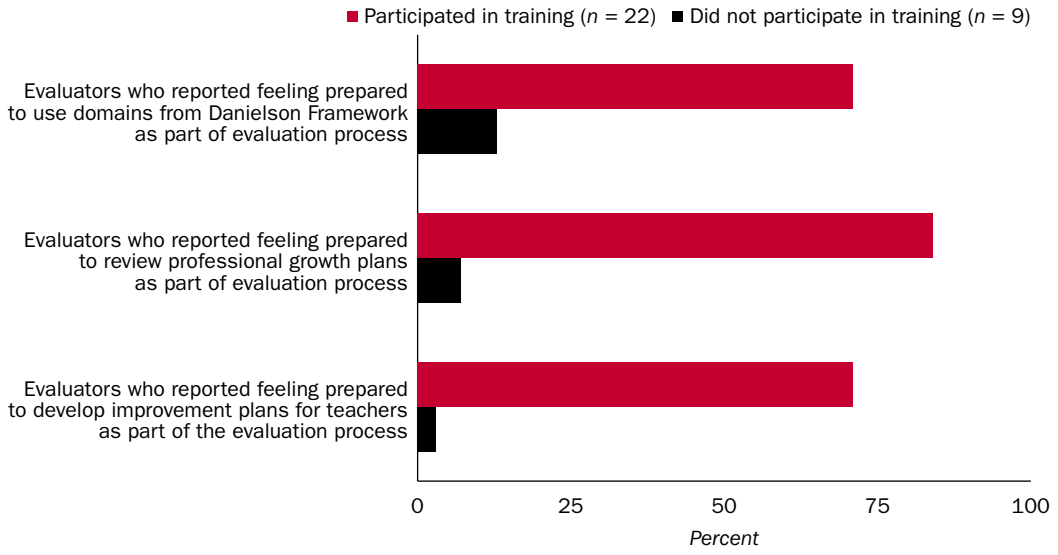
Many evaluators and teachers reported that the new evaluation systems took too long to complete and that there were too few evaluators to complete the required number of teacher evaluations. About 70 percent of evaluators and 62 percent of teachers reported that the system required too much time to implement. Interview data revealed more information about evaluators' and teachers' time limitations. From the principals' perspective, the system required considerable time to schedule and conduct classroom observations, walkthroughs, and conferences; compile the results from multiple measures for each teacher; and complete and maintain paperwork for all teachers. Teachers commented that it was cumbersome and time-consuming to complete paperwork and prepare for meetings with evaluators. Principals and teachers also described the inadequacy of resources to support the evaluations. For example, three interview respondents specifically commented on the need for additional support such as new technology or software to streamline the process of collecting data and completing the evaluations. One principal indicated that the school was looking for new software to calculate scores on evaluation or observation rubrics and that not all available software had the capability of doing this.

Training provided to evaluators

Training of evaluators was another factor related to implementation. While the state provided training support early in the summer before implementation in the following school year, especially for the Danielson Framework for Teaching,[3] classroom observations, and calibration of evaluations, not all evaluators participated in the training.

The state also offered ongoing training throughout the school year related to the Danielson Framework for Teaching and conducted classroom observations and walkthroughs more frequently than other features of the teacher evaluation systems (such as reviewing professional growth plans, teacher portfolios, classroom artifacts, teacher self-assessments, and student learning objectives). Evaluators who had participated in any trainings reported higher levels of preparedness to implement the features on which they had received training than evaluators who had not participated in trainings (figure 1).

Figure 1. Evaluator participation in training and preparation for implementation



Note: The sample included 31 evaluators.

Source: Authors' analysis of New Hampshire Department of Education teacher evaluator data, 2013.

In interviews, principals emphasized that they found the trainings they had received on how to use the Danielson Framework to be of high quality. In addition, they were pleased that they had had an opportunity to work with the evaluation rubrics before the New Hampshire Department of Education required their use. Several principals indicated that their schools were already using the Danielson Framework but that the calibration training was very useful. Principals also appreciated that the state-led training on the Danielson Framework had been scheduled a year in advance, which allowed principals to plan for conducting evaluations in the following school year. Some principals commented that they would have liked more opportunities throughout the implementation year for additional state-provided professional development and not just training related to the use of the Danielson Framework.

Introduction and development of student learning objectives

Principals reported that SLOs were more challenging to implement than other features of the new evaluation system (for example, conferences, walkthroughs, development of professional growth plans, and the like), particularly because student measures of learning had not previously been part of the evaluation process. Incorporating SLOs required considerable time, resources, and training to implement this new component of teacher evaluation, in part because there is little empirical research about the statistical properties of SLOs or their use in measuring student growth as a component of teacher evaluation (Gill, Bruch, & Booker, 2013).

Some 49 percent of evaluators reported feeling prepared to determine whether teachers had met their SLOs. As described earlier, a strong correspondence existed between initial training and evaluators feeling prepared to implement various features of the system—except for SLO training. Although 71 percent of evaluators indicated that they had received training on how to review SLOs, 53 percent of evaluators reported feeling prepared to review them. And while 62 percent of evaluators reported participating in training in how to determine whether teachers met SLOs, 49 percent reported feeling prepared to do so.

The SLOs represented perhaps the newest direction for the evaluation systems—while evaluators had experience with observations, conferencing, and in many cases some kind of walkthrough protocol, SLOs were an entirely new feature. In addition, although initial training provided by the New Hampshire Department of Education provided evaluators with some basic information about SLOs and how they functioned, the districts designed SLOs locally. As a result, there was considerable variation in how SLOs were designed, as well as variation in the capacity of school practitioners to support and measure progress in meeting the SLOs.

Interview data indicate that the training for SLOs seemed insufficient to fulfill evaluators' needs. Evaluators indicated that there were many questions and issues with implementation of SLOs, including how to identify appropriate sources of data, how to ensure rigor, and how to review and assess the SLOs themselves. Some principals reported that more specific and concrete examples of quality SLOs would be helpful, particularly for use with teachers in elementary school grades. Others indicated that they would like additional training on how to balance rigor with developing SLOs that were also attainable.

Stakeholder support from evaluators and teachers

A fourth factor related to implementation was the extent to which stakeholders supported the design and implementation of the new evaluation system. The majority of teachers and evaluators supported the evaluation system, with 83 percent of evaluators and 69 percent of teachers reporting that they think the evaluation system is fair (figure 2). Similarly, 74 percent of evaluators and 71 percent of teachers indicated that the teacher unions in their districts support the new evaluation systems. And 89 percent of

evaluators and 87 percent of teachers reported that teachers in schools are complying with the new evaluation requirements.

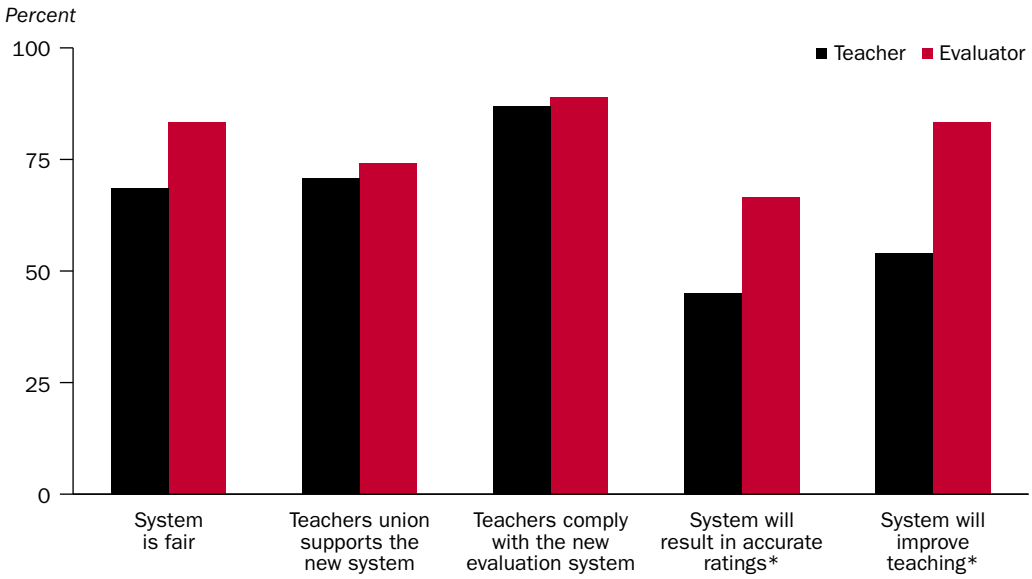
Teachers and evaluators did not agree in their perceptions of the long-term benefits of the new evaluation systems: 67 percent of evaluators and 45 percent of teachers believed that the new evaluation system would result in accurate ratings of teachers.[4] Similarly, 83 percent of evaluators and 54 percent of teachers thought that the new system would improve teaching (see figure 2).[5]

Teacher support for the new evaluation system seems to be related to implementation fidelity (defined as the percentage of teachers in a district that reported being evaluated on the required features of the system). The three districts with the highest average fidelity also had the highest means on the survey for fairness/compliance and support of desired implementation outcomes.[6] However, it is unknown whether higher stakeholder support facilitated higher implementation fidelity or whether higher implementation fidelity led to higher stakeholder support.

Teachers' perceptions of professional climate in their schools

The fifth factor related to implementation was the professional climate of schools. The teacher survey used in the study included items designed to measure perceptions of professional climate. These items were adapted from the Chicago Consortium for School Research (2012) survey on school climate. It included constructs of leadership, teacher influence, and trust among peers and leaders.[7] Schools with a more favorable climate—for example, schools in which teacher trust in administrators and influence in school-level decisions was high—had greater implementation fidelity. For more information about implementation fidelity, see the companion report, Riordan, Lacireno-Paquet, Shakman, Bocala, & Chang (2015).

Figure 2. Teacher and evaluator perceptions of new teacher evaluation systems



* Differences between teachers and evaluators are statistically significant at the $p < .05$ level.

Source: Authors' analysis of New Hampshire Department of Education teacher and evaluator survey data, 2013.

Policy implications for state and district design and implementation

The factors identified in this brief raise several policy implications for states and districts implementing educator evaluation systems:

- *Assess and address capacity issues for evaluators.* This includes assessing evaluators' time and the resources they have to complete the evaluations. As suggested by some teacher and principal interviewees, specific strategies to address limited capacity might include reducing the number of observations required for more experienced teachers who have been rated proficient on previous evaluations or reducing the frequency of walkthroughs.
- *Provide adequate planning time or introduce components incrementally to support the implementation of new and complex initiatives.* The 15 SIG schools had a school year to redesign a system for evaluation aligned with the state framework and to develop an implementation plan. Several districts phased in parts of the plan to ease the transition, such as not measuring teachers' performance on all elements of the Danielson Framework for Teaching in the first year or phasing in the requirements related to the types of SLOs.
- *Allow for adequate early and ongoing training on the system.* This is important for building support for the system and assuring that participants are prepared to participate in evaluations. State officials provided training for evaluators early in the summer, before new evaluation procedures were enacted, including training on the Danielson Framework for Teaching, classroom observations, and calibration. While this initial support was valuable, interview respondents also indicated that ongoing training was critical to the success of their implementation.
- *Training for SLOs requires special attention.* SLOs were one of the most challenging aspects of implementation in New Hampshire. This suggests that state or district officials should provide repeated opportunities for practitioners to meet with measurement experts and build their understanding of implementing and measuring SLOs, while considering evaluators' and districts' capacity to implement these new measures.
- *Engage stakeholders.* Stakeholder support is important to implementation and may be enhanced by engaging stakeholders in the design process. In New Hampshire the original framework for evaluation was developed by a broad stakeholder group. In turn, each SIG district had a year to develop its specific plan with a local stakeholder group made up of teachers and administrators. The majority of administrators and teachers indicated that they supported the new evaluation system and their engagement in the design process may be related to this generally widespread support.
- *Foster a positive professional climate.* This involves including teachers in informal and formal decisionmaking. The findings of this study provide some preliminary understanding that professional climate may play a role in the implementation of new teacher evaluation systems. There appears to be a more favorable climate among districts with higher fidelity, especially as it relates to teacher influence in school decisions and trust among teachers and administrators.

Notes

The authors wish to thank members of the Northeast Educator Evaluation Research Alliance and the New Hampshire Department of Education for their support in conceptualizing this study. In particular, the authors appreciate the contributions of New Hampshire Department of Education Commissioner Virginia Barry, Kathryn "Joey" Nichol, and Karen Soule to the design, data collection, and discussion of the dissemination of findings as well as to earlier drafts of this report.

1. School improvement grants are federal funds distributed by states to local education agencies to provide financial assistance for school improvement activities. States that receive grants must give priority to the lowest achieving schools that also demonstrate the greatest need for the funds and a strong commitment to using the funding to meet school improvement goals.
2. Evaluators include any administrative staff responsible for conducting teacher evaluations. In some schools only principals were evaluators, and in other schools other administrators such as assistant principals shared this responsibility.
3. The Danielson Framework for Teaching is a set of 22 components of instruction (for example, setting instructional outcomes) that are aligned to the Interstate Teacher Assessment and Support Consortium standards. The components are divided across four domains of teaching responsibility: planning and preparation, classroom environment, instruction, and professional responsibilities. All SIG schools in the study employed the Danielson Framework and its domains, although components and weighting varied.
4. This difference was statistically significant ($p < .05$).
5. This difference was also statistically significant ($p < .05$).
6. See Riordan et al. (2015).
7. Teachers were asked to indicate the extent to which they agreed or disagreed with statements related to these constructs (for example, "I trust the principal.").

References

Chicago Consortium for School Research. (2012). *Survey documentation and measure information and statistics*. Retrieved March 28, 2014, from https://ccsr.uchicago.edu/surveys/documentation.

Gill, B., Bruch, J., & Booker, K. (2013). *Using alternative student growth measures for evaluating teacher performance: What the literature says* (REL 2013–002). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. http://eric.ed.gov/?id=ED544205.

National Council on Teacher Quality. (2012). *State of the states 2012: Teacher effectiveness policies*. Washington, DC: Author. http://eric.ed.gov/?id=ED536371

Riordan, J., Lacireno-Paquet, N., Shakman, K., Bocala, C., & Chang, Q. (2015). *Redesigning teacher evaluation: Lessons from a pilot implementation* (REL 2015–030). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2015030.pdf.

REL 2016–101

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

November 2015

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0009 by Regional Educational Laboratory Northeast & Islands administered by Education Development Center. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education; nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Riordan, J., Lacireno-Paquet, N., Shakman, K., Bocala, C., Chang, Q. (2015). *Stated Briefly: Redesigning teacher evaluations: Lessons from a pilot implementation* (REL 2016–101). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

The Regional Educational Laboratory Program produces 7 types of reports