



MINNESOTA

PRINCIPAL EVALUATION

MODEL PILOT STUDY

Commissioned by the Minnesota Department of Education

SPONSORED BY THE BUSH FOUNDATION | CONDUCTED BY FHI 360 | APRIL 2014

fhi
360
THE SCIENCE OF
IMPROVING LIVES

Table of Contents

Introduction	3
Data Collection and Respondents' Characteristics	5
Report Purpose	7
Part I. Overview of the Principal Evaluation Example Model: Components and Implementation Process	9
Three Components of the State Example Model	12
Five Performance Measures	13
Performance Levels	13
Rubrics	13
Model Implementation Process—Eight Steps	15
Part II. Value and Fairness of the Example Model	17
1. How much time did principals and evaluators invest in the implementation of the example model or their adaptation of the model?	21
2. Did evaluators and principals think their time was well spent?	25
3. What components of the model did the principals and evaluator think were most valuable?	27
4. On which performance measures and other topics did the principals receive the most feedback?	31
a. Evaluators' Feedback to Principals	31
b. Principals' Perceptions of their Evaluators' Feedback	33
5. How did feedback to principals vary by district characteristics?	35

6. What is the value added of participation in principal evaluation for principal’s professional practice?	37
7. What was the perception of the capacity of various data sources to measure principal performance accurately and fairly?	39
a. Principals’ Mid- and End-of-the-year Responses	42
8. What skills do evaluators need in order to provide effective feedback to principals?	45
Summary comments	47
Part III. Implementation of the Example Model at the School and District Levels	49
1. How well did principals and evaluators understand the example model?	52
2. What did the implementation of the example model look like at the district level and how did it vary?	53
a. Self-Assessment and Goal-Setting	53
b. Preparation for the Mid-year Conference	55
c. Preparation for the End-of-year Conference—Variations between Principals and Evaluators	61
d. End-of-year Conferences—Variations between Principals and Evaluators	65
e. Evidence Used for the Quantitative Rating by the Case Study Evaluators	70
f. How Case Study Evaluators Calculated the Final Rating	72
g. Perceptions of the Final Ratings	72
h. Case Study Evaluators/ Reflections on the End-of-year Conference Process	77
3. What challenges were faced in implementing principal evaluation?	78
a. Challenges Reported by the Case Study Evaluators	79
b. Challenges Reported by the Case Study Principals	80
4. What Conditions Support Successful Implementation of Principal Evaluation?	82
a. State Conditions	83
b. District Conditions	83
Part IV. Recommendations	87
1. Recommendations for State Actions to Strengthen Implementation	90
2. Recommendations for District Actions to Strengthen Implementation	91
3. Recommendations for State Actions to Strengthen the Design of the Example Model	92
Questions for Further Investigation	95
In Summary	96
Appendix A: Details of Recommendations to Strengthen the Implementation of the Model	97
1. Recommendations for State Action to Strengthen Implementation	98
2. Recommendations for District Action to Strengthen Implementation	101

Introduction

IN MAY 2011, the Minnesota State Legislature mandated that all districts evaluate principals annually. To support the evaluation process, the guidelines in the statute called for the development of an evaluation model that would improve the instructional leadership and organizational skills of principals, as well as strengthen their capacities “in the areas of instruction, supervision, evaluation, and teacher development.”¹ Legislators also wanted a mechanism to hold principals more accountable for school outcomes, although the primary and explicit emphasis of the model was to be on principals’ professional growth.

The Principal Evaluation Working Group,² appointed by the Minnesota Department of Education (MDE) Commissioner, included representatives from the Minnesota Association of Secondary School Principals and the Minnesota Association of Elementary School Principals. It submitted the Minnesota Principal Evaluation Model (hereafter referred to as the example model) to the legislature in February 2012. During the 2012–2013 school year, MDE piloted principal evaluation in 17 districts, selected in part to vary by size, location, student demographics, and achievement outcomes.

Districts had flexibility in how they implemented principal evaluation; most used at least some parts of the example model or refined it to their context. It is important to note that this report focuses both on the piloting of an annual principal evaluation process in the 17 districts and on feedback and issues specific to the example model they were given for

¹ Minnesota Statute 123B.147, 2011.

² The working group included principals, superintendents, business leaders, directors of principal and administrator associations, parents, school board members, teachers, university professors, and the state licensing board.

guidance. The report aims to help inform improvements to principal evaluation overall and specifically to the example model to support its evolution to an even more valuable resource to districts. Thus, the report uses the term “example model” only when referring to the model designed by the working group; in other cases, we discuss principal evaluation or the principal evaluation process.

The example model recommended by the working group consists of three components:

- 1) annual principal evaluation by the superintendent using a performance-based system;
- 2) use of proficiency and longitudinal data on student academic growth from sources (such as the Minnesota Comprehensive Assessment [MCA], Northwest Educational Associates Measures of Academic Progress [MAP]) and/or graduation rates that incorporate district achievement goals and targets; and
- 3) feedback from other school community stakeholders, including students, parents, colleagues, and staff.

This report presents findings and recommendations from a study of the implementation of principal evaluation in the 17 pilot districts during the 2012–2013 school year.³ The findings draw from data collected through confidential online surveys completed by principals and evaluators in the 17 pilot districts and interviews with principals and evaluators in four “case study” districts selected to help provide context for the survey data.

Data collected in March 2013 corresponded to the completion of the mid-year conference; it also included the goal-setting process.⁴ Data collected from September to October 2013 corresponded to the completion of the end-of-year conference and included reactions to the mid-year conferences.

³ The evaluation process extended through October 2013 because summative reviews of principals in some districts were postponed until the student achievement data were received that fall.

⁴ A report on the findings from this round of data collection, *Voices from the Field*, covers study findings through the mid-year conferences can be found at the MDE Web site.

All information gathered through the surveys and interviews is being used solely for analytic purposes, and the identity of respondents and case study districts is confidential.

**Data Collection
and Respondents’
Characteristics**

Survey Data

Principals and evaluators in pilot districts responded to two surveys about the 2012–2013 pilot year. The **mid-year survey** focused on their understanding of the example model and the scope of work needed to implement the evaluation process. The mid-year survey also asked respondents about the training they received in the example model, their self-assessment and goal-setting conference, and their preparation for and participation in the mid-year conference. The **end-of-year survey**, on the other hand, sought to collect information on activities principals and evaluators engaged in while preparing for the end-of-year conference, and the end-of-year conference itself. The final survey also asked respondents to reflect on the value of the principal evaluation process, challenges encountered throughout the year, and recommendations on how to improve the model. As mentioned earlier, because many districts waited to conduct the summative reviews until the release of state student achievement data in the fall of 2013, the survey timeline was extended into the 2013–2014 school year.

About the Survey Respondents

Over the course of the pilot study’s two surveys, a total of 102 principal and 25 evaluator⁵ surveys were collected. The response rates among both principals and evaluators were similar across each data collection period: 51% of all principals in the pilot districts completed mid-year and end-of-year surveys, and 61% and 70% of evaluators in the pilot districts participated in the mid-year and end-of-year surveys, respectively.

Although the respondents of the principal surveys primarily included school principals (over 90%), four respondents were assistant principals or deans. In both surveys, 45% of respondents work in Pre-K, kindergarten, or elementary schools; another one-fourth works at the high school level. The remaining approximately 30% of respondents work in schools with other configurations such as middle schools, elementary/middle, middle/high, and alternative schools.

Interview Data from Case Study Districts

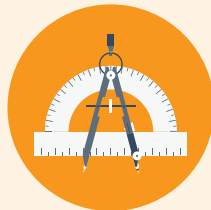
In addition to the principal and evaluator surveys, four case study districts were selected in consultation with MDE. The goal of the case study analysis was to gather deeper insights about how well the pilot was implemented within a district context, the types of challenges encountered in using the principal evaluation process and ways they were addressed, and potential changes or promising practices to be recommended as state-wide principal evaluation moves forward.

⁵In the case studies the evaluator was typically the superintendent but could also include other district staff. In this report we use “superintendent” and “evaluator” interchangeably.

EXAMPLE MODEL



1. Annual principal evaluation by the superintendent using a performance-based system



2. Use of proficiency and longitudinal data on student academic growth from sources



3. Feedback from other school community stakeholders, including students, parents, colleagues, and staff

Interviews in these districts occurred concurrently with administration of the surveys.

The selected districts varied according to district size as measured by the total number of students enrolled, location of the district (urban, rural, suburban), student achievement, superintendent tenure, and racial/ethnic and socioeconomic diversity of the student body. In each of the selected districts, interviews were conducted with two principals from schools with different grade levels and their evaluator. The research team conducted both in-person and phone interviews and always interviewed principals separately from evaluators.

About the Case Study Participants

Researchers interviewed eight principals in the spring of 2012 and fall of 2013. One district dropped out of the pilot so another district was added. Eventually four evaluators and eight principals were interviewed after the mid-year conference and after the end-of-year conference. Each interview, lasting between 60 and 90 minutes, generally took place during work hours in the respondent's office. Each interview was then transcribed and analyzed using a narrative analysis technique.

Report Purpose

This mixture of design methods allowed the research team to collect rich data on the implementation of nearly every element of the example model. It provided principals and evaluators with an opportunity to share what they felt worked well and to reflect on the challenges experienced while working with the model and, in most cases, adapting it to their contexts as needed. This report paints a genuine picture of the principals' and evaluators' experiences implementing principal evaluation. It uses their feedback to share helpful experiences and findings and to develop specific recommendations on how the model and its implementation can be improved.

The report describes the principal evaluation implementation process at the district level, using both mid-year and end-of-year survey and interview data, although it primarily focuses on the end-of-year processes.⁶ It also offers recommendations to the MDE and its working group on how to strengthen the example model. We believe the model has and can continue to provide powerful guidance to all Minnesota districts, which must now (as of the 2013–2014 school year) be implementing principal evaluation practices based on the state statutory requirements. The knowledge gained by piloting principal evaluation provides valuable insights to help the districts address the state's requirements.

⁶ The report on the first phase of the pilot, *Voices from the Field*, covers study findings through the mid-year conferences, and was submitted in May 2013. It can be found at the MDE Web site.

PART I.

OVERVIEW OF THE PRINCIPAL EVALUATION EXAMPLE MODEL

COMPONENTS AND IMPLEMENTATION PROCESS

PRIOR TO THE new state principal evaluation model, Minnesota principals were required to have a valid license in supervision and administration and were assigned by the school board to lead a school upon the recommendation of the district superintendent. Their duties included providing “administrative, supervisory, and instructional leadership services, under the supervision of the superintendent of schools of the district and in accordance with the policies, rules, and regulations of the board of education, for the planning, management, operation, and evaluation of the education program of the building or buildings to which the principal is assigned.”⁷

In the case studies we learned of a variety of prior principal evaluation practices that ranged from very informal individual meetings between principals and the superintendent (often on a five-year cycle) to annual evaluations that were similar to the example model. The example model adds more structure to the processes the districts used for principal evaluations while also allowing leeway for local flexibility and some adaptations.

⁷ Minnesota Statute 123.B.147 Subdivision 3., 2010.

THREE COMPONENTS OF THE STATE EXAMPLE MODEL

The recommended state model includes three major components based on the statute.



COMPONENT ONE is the supervisory evaluation conducted by the superintendent or a designee who has been appropriately trained. This is to include on-the-job observations and information from previous evaluations. The evaluator is responsible for conducting on-the-job observations as part of an annual evaluation of each principal. The evaluation must be consistent with the job description, a district's long-term plans and goals, and the principal's own professional multiyear growth plans and goals. Performance measures with more detailed indicators were adopted by the working group to be used as the basis for reviewing the principal's work. They include rubrics that support completing this portion of the evaluation.



COMPONENT TWO is school-level performance data tied to established district goals related to student outcomes. This satisfies the component in the statute specifying the use of longitudinal data on student academic growth as an evaluation component that incorporates district achievement goals and targets and counts for 35% of a principal's final rating. The state example model incorporates performance data through a goal-setting process. The principal is asked to set specific goals related to student achievement and school performance and is then evaluated on how well he or she attained those goals during the evaluation period. The goal setting is done in collaboration with the evaluator.



COMPONENT THREE is the use of surveys and other feedback from teachers and community stakeholders. The surveys or feedback are to be designed to help identify a principal's effectiveness, leadership skills and processes, and strengths and weaknesses in exercising leadership in pursuit of school success. The state model requires that feedback from stakeholders on a principal's performance be part of the summative evaluation and that, at a minimum, it includes feedback from teachers. A district may choose, but is not required, to gather feedback from other stakeholders. The local district determines the type of instrument used to gather this feedback.

When combined, these three areas determine a principal's summative evaluation and performance rating.

Five Performance Measures To establish a shared vision and common expectations for quality leadership in Minnesota schools, five performance measures were developed as part of the example evaluation model.⁸ These measures form the foundation for all components of the example model. Performance measures reflect what effective principals should know and be able to do. For the example model, principals should be able to effectively:

1. Establish a vision and mission focused on shared goals and high expectations
2. Provide instructional leadership for high student academic performance
3. Manage human resources for quality instruction and professional growth
4. Build professional and ethical relationships through collaboration and effective communication
5. Strategically manage resources for systemic performance accountability

Performance Levels The Minnesota state model contains four performance categories for each indicator: Distinguished, Accomplished, Satisfactory, and Unsatisfactory. In addition, a designation of Developing may be assigned to a principal undertaking a new assignment or a principal in a probationary period. It also may be used if a significant change has occurred in district goals, curricula, leadership, or strategic vision. This designation may apply to any of the four performance designations.



Rubrics Rubrics for each indicator, developed by a committee of the working group, are included as part of the model. The rubrics provide detailed guidance for making fair, accurate, and consistent judgments about performance. Using the rubric to establish a common understanding of expectations contributes to quality assurance in the process for both the evaluator and the principal. The rubrics also can serve as the basis for an ongoing dialogue about results, as a tool for systemic feedback, and as guidance for developing growth plans.

A sample rubric for one indicator is on the next page.⁹ The summative rating, described on the next page, reflects a compilation of the rubric ratings for each indicator.

⁸ These measures are based on research on the skills principals need to raise student achievement. See for example, *The School Principal As Leader: Guiding Schools to Better Teaching and Learning*. The Wallace Foundation, January 2013.

⁹ From *Measuring Principal Performance in Minnesota: Performance Measure and Indicators Rubric*, Minnesota Department of Education, January 30, 2012

Performance Measure #1: Establishes a Mission and Vision Focused on Shared Goals and High Expectations

Indicator 1B: Articulates a Vision and Develops Implementation Strategies for Change that Result in Measurable Achievement Gains for All Students and a Closing of the Achievement Gaps

DISTINGUISHED (4)	<ul style="list-style-type: none"> Principal initiates and communicates changes to vision, mission, and goals based on data to improve school performance and establishes a school culture of high expectations that incorporates collaborative decision-making processes to achieve measurable goals and close the achievement gap for all students. Principal identifies and communicates actions, roles, responsibilities, timelines, and decision-making processes to appropriate stakeholder groups. Principal builds staff ownership, efficacy, and empowerment in the goal areas and assigns teachers to focus on goals effecting students school-wide as well as specific goals appropriate to their assignment areas. 	<ul style="list-style-type: none"> Principal designs, develops, and implements ongoing cycles of data collection, analysis, professional development, implementation, and reflection to monitor progress toward goals and to promote change for continuous improvement. Principal adapts his or her leadership style to meet the needs of specific situations, models flexibility, and builds a sense of efficacy and empowerment among staff to collectively meet challenges, increase student achievement, and improve teacher effectiveness.
ACCOMPLISHED (3)	<ul style="list-style-type: none"> Principal communicates and models a commitment to the vision, mission, and goals and promotes a school culture of high expectations that incorporates collaborative decision-making processes to achieve measurable goals and close the achievement gaps for all students. Principal connects actions and communications, and includes all stakeholders in decision-making processes. Principal engages a diverse group of stakeholders and the support of the central office to implement changes needed to improve learning. 	<ul style="list-style-type: none"> Principal creates a process to gather data to monitor, track, and review progress toward goals and routinely and systematically communicates impacts and progress to stakeholders. Principal monitors the change process and addresses factors that will increase staff motivation and installs practices that promote persistence and well-being.
SATISFACTORY (2)	<ul style="list-style-type: none"> Principal communicates a commitment to the vision, mission, and goals and promotes a school culture of high expectations, without embedded processes to support vision, mission, and goals. Principal communicates the decision-making process to key stakeholders in the development of specific and measurable achievement goals. Principal identifies changes needed to improve student learning and engages staff through effective communication and strategies. 	<ul style="list-style-type: none"> Principal remains focused on the school vision and mission as he or she identifies goals and uses data to monitor progress toward goals. Principal understands how adults learn and uses this understanding to plan professional development and support the continuous learning process.
UNSATISFACTORY (1)	<ul style="list-style-type: none"> Principal actions lack emphasis on learning and/or high and measurable expectations in implementing vision, mission, and goals. Principal fails to connect actions, communications, and decision making to the organization's vision, mission, and goals. Principal fails to monitor impact of change strategies and does not track progress against the plan in order to adjust strategies as needed. 	<ul style="list-style-type: none"> Principal fails to use relevant data to monitor progress toward goals. Principal is unable to constructively respond to challenges, and does not appear to understand the importance of building a sense of efficacy, empowerment, and well-being among staff.

Model Implementation Process

Eight Steps

There are eight steps to implementing the model, usually starting just before the academic year begins and running through October of the following academic year, after the state assessment data are available. To support this process, the MDE created recording forms that capture most steps of the process. For example, the principals used the performance measure rubrics to record their self-assessment. Another form was used to record the results of the self-assessment and professional goal setting. The evaluator's evaluation and the summative performance measure ratings were recorded on another form.¹⁰

1. Orientation On The Process The evaluator provides the principal with a complete set of materials outlining the evaluation process and (as appropriate) with his or her summary evaluation form from the last performance review.

2. Pre-Planning by the Principal The principal reviews student achievement data from the previous school year, the prior year's evaluation results, and other relevant data (for example, artifacts, survey results, teacher retention data). The principal uses these data sources to

- Reflect on his/her own leadership practice.
- Self-assess current performance on the performance measures and indicators.
- Identify priorities for professional growth and performance goals.

The principal submits his or her self-assessment to the evaluator, who reviews it along with school results and other relevant data, then sets a meeting to review and finalize priorities for the school year—the goal-setting conference.

3. Goal-setting Conference The evaluator and principal meet to discuss, as applicable, the last performance evaluation, self-assessment, priorities for professional growth, and preliminary performance goals. At the conclusion of the conference, the evaluator and principal come to agreement on performance goals (in the absence of agreement, the evaluator will make the final determination about the goals). The conference is an opportunity to review the evaluation process overall and the processes to be used by the evaluator and principal to collect evidence about the principal's performance. If the evaluator has significant concerns regarding the principal's performance and intends to use a process that is targeted at remediation, this should be communicated clearly at the goal-setting conference.

4. Evidence Collection The principal collects evidence agreed upon in the goal-setting conference. This may include data for each performance measure included in the review; feedback from parents, students, and the school community; documentation of the professional development achieved during the year; evidence of student growth and success measured over time; and other data to document achievement of performance goals. In addition, the evaluator will directly observe principal practice, interact with teachers and other members of the school community, and gather additional evidence to support the review. The evaluator should provide timely feedback on his or her observations, as appropriate. (This step is repeated in the spring after step 5).

5. Mid-year Conference The evaluator and principal hold a mid-year formative conference, with explicit discussion of progress against growth goals and all performance measures, as well as interim student and teacher performance data.

6. End-of-year Summative Conference The principal submits any evidence agreed to in the goal-setting and mid-year conferences. The evaluator reviews this evidence and all other evidence gathered by him or herself and assembles a preliminary summary rating of the principal. The evaluator and principal hold an end-of-year conference to review and discuss accomplishment of growth goals, the preliminary summary rating, and interim student and teacher performance data.

7. Summative Performance and Final Rating Following the conference, the evaluator finalizes the summary rating and generates a summary report of the evaluation, which is signed by both parties. A performance level is assigned using one of the categories mentioned above (for example, Distinguished–Unsatisfactory)

8. Growth Plans A professional growth plan is developed from the summary report, consistent with the assigned performance rating. Growth plans for the coming year (and sometimes for multiple years) are signed by both parties. Depending on the principal's summative rating, the principal will have more or less authority to direct his or her growth plan.

4 – Distinguished: A self-directed growth plan; eligible for additional leadership roles and responsibilities; encouraged to assume role of mentor or coach.

3 – Accomplished: A self-directed growth plan.

2 – Satisfactory: One-year jointly designed growth plan.

1 – Unsatisfactory: One-year directed improvement plan stemming from unsatisfactory or troubling performance items; generated by the evaluator and specifically identifies areas for improvement.¹¹

Developing: In addition to the four ratings above, a principal may be designated as Developing if a limited number of items targeted for improvement are added to the growth plan by the evaluator under the criteria listed for this designation. This should be viewed as a transition plan that is designed to address new knowledge and skills required by a change in position or changes initiated at the district level.

¹⁰ These forms are available in Principal State Example Model: Pilot Resources and Forms, Minnesota Department of Education. <http://education.state.mn.us/MDE/EdExc/EducEval/PrincEval/index.html>

¹¹ For Unsatisfactory, failure to remedy or improve results in disciplinary action per Minnesota Statutes, section 122A.40 or 122A.41 or local district policy.

PART II.

VALUE AND FAIRNESS OF THE EXAMPLE MODEL

“This is by far, for me, the most professional work that I’ve done for myself.”

–Principal interview, Fall 2013

OVERALL, AS THIS section will examine, the study findings strongly indicate that pilot participants believe the example model to be both fair and valuable in supporting principal professional growth. Evaluators and principals agreed that the pilot met their expectations of improving school leaders' professional practice and school performance.

The findings below address the following eight core study questions:

1. How much time did principals and evaluators invest in the implementation of the example model or their adaptation of the model?
2. Did evaluators and principals think their time was well spent?
3. What components of the model did the principals think were most valuable?
4. On which performance measures and other topics did the principals receive the most feedback?
5. How did feedback to principals vary by district characteristics?
6. What is the value added of participation in principal evaluation for principal's professional practice?
7. What was the perception of the capacity of various data sources to measure principal performance accurately and fairly?

The findings reflect a mix of (a) factual issues such as time spent and variation by district characteristics and (b) perceptions of the evaluation process, such as most valued components.

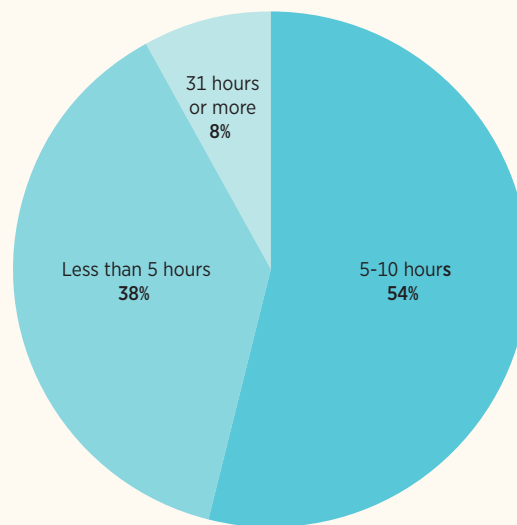
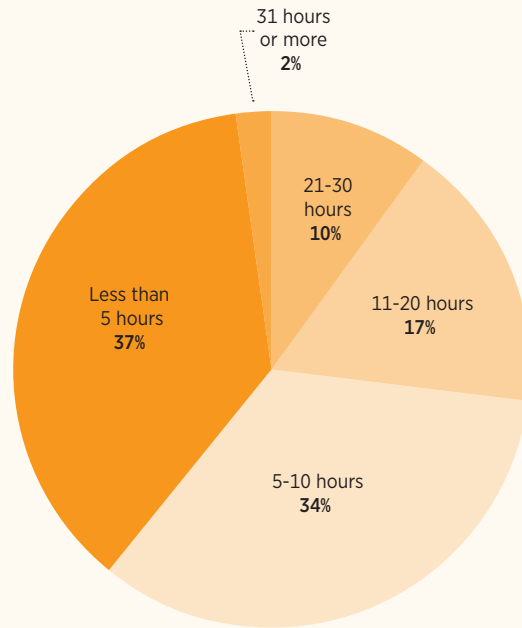


FIGURE 1.

Thinking about your implementation of the principal evaluation process during 2012-13, approximately how many hours have you spent on evaluation-related activities on a typical week?

SOURCE: Principal Survey

FIGURE 2.

Thinking about your implementation of the principal evaluation process during 2012-13, approximately how many hours have you spent on evaluation-related activities on a typical week?

SOURCE: Evaluator Survey

1. How much time did principals and evaluators invest in the implementation of the example model or their adaptation of the model?

Survey Results

In the end-of-year survey, we asked both evaluators and principals about the time they invested in the implementation of their district's evaluation model. We learned that evaluators primarily spent time preparing for the summative conference, conducting the conference, and calculating each principal's rating. Principals also worked on preparing for the summative conference, but their focus was on evidence collection and "working their action plans" to achieve their goals.

In the end-of-year case study interviews, we asked questions that teased out the specifics of how evaluators and principals spent their time. These time investments included revising and aligning the performance measure rubrics to their district's leadership and school improvement strategies; linking principals to district support personnel to help them with goal setting and/or to provide supports to achieve their goals; using district administrative meetings to deepen principals' understanding of the performance measures; and setting up spreadsheets to organize principals evidence of attaining their goals and to facilitate the rating calculation.

In the final survey, 92% of the evaluators and 71% of the principals reported that they spent 10 hours or less weekly on principal evaluation-related activities during the 2012–2013 school year (see Figures 1 and 2).

The survey responses indicated that evaluators invested most of their time on the implementation of the end-of-year conference, giving feedback as well as assigning performance ratings and completing summative evaluation forms (see Figure 3 below).

From the interviews in the four case study districts, we learned that evaluators invested significant effort into making evaluation implementation a priority. Specifically, one district spent several administrative meetings with principals to review and revise the rubric descriptors, making them more specific to the levels of performance that aligned with expected practices in their district. Another district organized a panel of principals and administrators to do the same thing, which gave credibility and fairness to the assignment of each principal's performance rating in the rubric.

FIGURE 3.

The Percent of Evaluators who Rated Different Elements of the Example Model as Most Time-Consuming (N=13)



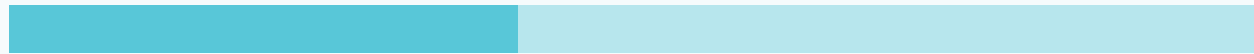
15.4% Supporting Principals During their Self-Assessment



15.4% Preparation for and Supporting Principals in their Preparation for the Mid-Year Conference (i.e. evidence collection and review)



23.1% Preparation for and Supporting Principals in their Preparation for the End-of-Year Conference (i.e. evidence collection and review)



38.5% Supporting Principals in their Professional and School Improvement Goal-Setting



46.2% Mid-Year Conference and Feedback Giving



76.9% End-of-Year Conference and Feedback Giving



76.9% Completing Required Forms for Summative Evaluation and Assigning Performance Ratings

Several of the evaluators also talked about engaging other district personnel to support both them and principals throughout the year. In one district, for instance, district personnel helped principals shape their goals and provided support to principals throughout the year on implementing their action plans. In another case study district, an equity specialist helped principals set goals, familiarized them with research, and helped them design and implement plans to close the achievement gap.

Last year, we created a new department called Evaluation and Assessment and so [that director] helps them with their goals. [She/he] has to sign off on their goals—their academic goals and their stakeholder goals.

To help principals better understand their evaluation systems, one district dedicated one of its bi-weekly meetings with principals to review the practical application of performance measures and what they would look like in practice in the district. One evaluator reported spending funds for services to provide, conduct, and report stakeholder surveys and track these and other results. Finally, another district used time and funds to set up spreadsheets to help organize and analyze the principal evaluation data as well as to compute principal ratings.

...we came up with just a brief Google spreadsheet and basically what they [principals] had to do then was put their goals in there and then put their actual [student test] scores from last year and then what they were from this year and what their goal was and then the formulas did the work. So then it's just automatically set if they met their goal or not.

Principals in the case study districts, on the other hand, reported that their time and resource investments were most often related to setting their goals; taking actions in their schools to meet their goals; collecting evidence to document their goal attainment; engaging in activities that would support their professional learning; and participating in the ongoing evaluation process and evaluation conferences.

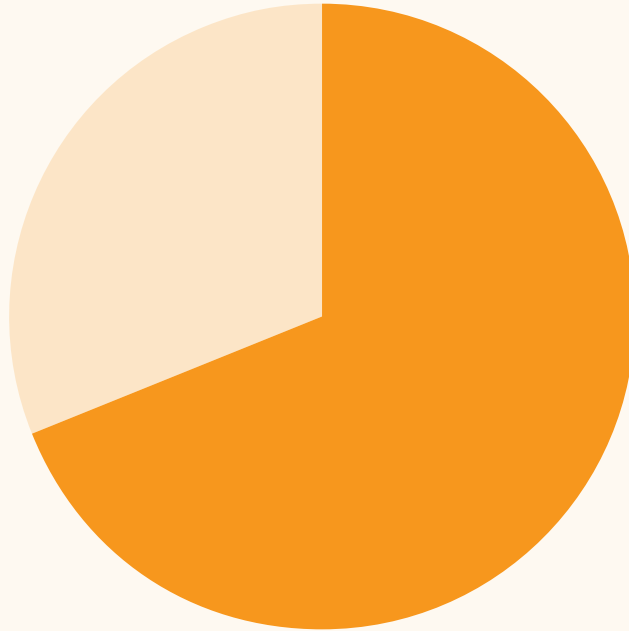
What were the implementation costs?

In the case study interviews, we sought more specifics on where the additional district resources were spent. As mentioned above, one district contracted with a vendor for the stakeholder surveys. Another purchased software to assist principals in providing feedback to teachers on their instructional practices, hired short-term coaches to help develop principals' skills, and subsidized the cost of principals' professional development (for example, attending workshops related to professional and school improvement goals on topics such as leading professional learning communities (PLCs),¹² community engagement, and parent involvement).

For the most part, principals in the case study districts worked within their school budgets to accomplish their goals, but in some cases the districts provided school leaders with additional resources to support their professional and school goal achievements. These included purchasing or expanding successful programs to extend student achievement results in other grades, purchasing a classroom observation tool, or sending principals to workshops to learn new skills and strategies.

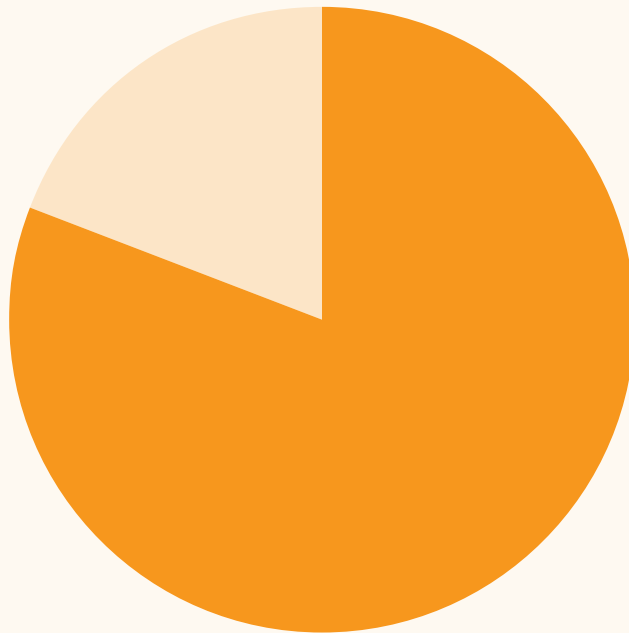
¹² A professional learning community, or PLC, is a group of educators that meets regularly, shares expertise, and works collaboratively to improve teaching skills and the academic performance of students.

PRINCIPLE SATISFACTION
regarding the outcome of the evaluation process



69%

Mid-Year Survey



80.9%

End-of-Year Survey

2. Did evaluators and principals think their time was well spent?

Responses to the final survey showed that participation in the pilot implementation of principal evaluation met both evaluators' and principals' expectations for improving school leaders' professional practice and school performance. Principals' satisfaction with the outcomes of the evaluation process improved in the course of the year. On the mid-year survey about 69% of principals agreed that the outcomes of their evaluation would be worth the invested time and effort; by the end-of-year survey, 80.9% of the principals felt that the model would help them improve their professional practice in the future.

Despite the considerable time and financial investment that the implementation of principal evaluation required, most evaluators and principals in the case study districts also agreed that the outcomes of the model implementation justified the costs. As one evaluator summed it up:

I think our feeling was that it was time well spent. Principals also felt that it was giving them good and honest feedback about what they were doing, and they felt that they were getting something out of the reflection process.

TABLE 1. Principals and Evaluators who Assessed Various Elements of the Model as Valuable

Elements of the Principal Evaluation Process	Percent of Principals (N=48)	Percent of Evaluators (N=13)
Self-Assessment	87.8%	100.0%
Goal-Setting	85.3%	100.0%
Mid-Year Review Conference	69.2%	100.0%
End-of-Year Conference	79.5%	100.0%
Review of Student Achievement Data	70.7%	92.3%
Teacher Feedback	65.0%	76.9%
Evidence Collection	61.9%	69.2%
Evaluator's Feedback Based on Observations and Interactions with Members of The School Community	67.5%	69.2%
Student Feedback	46.4%	23.1%
Parent Feedback	47.6%	23.1%
Feedback from Other School Community Members	50.0%	0.0%

3. What components of the model did the principals and evaluator think were most valuable?

Both evaluators and principals singled out these five elements as the most valuable to their professional growth: (1) self-assessment, (2) goal-setting conference, (3) mid-year and (4) end-of-year conferences, and (5) student achievement data review (see Table 1).

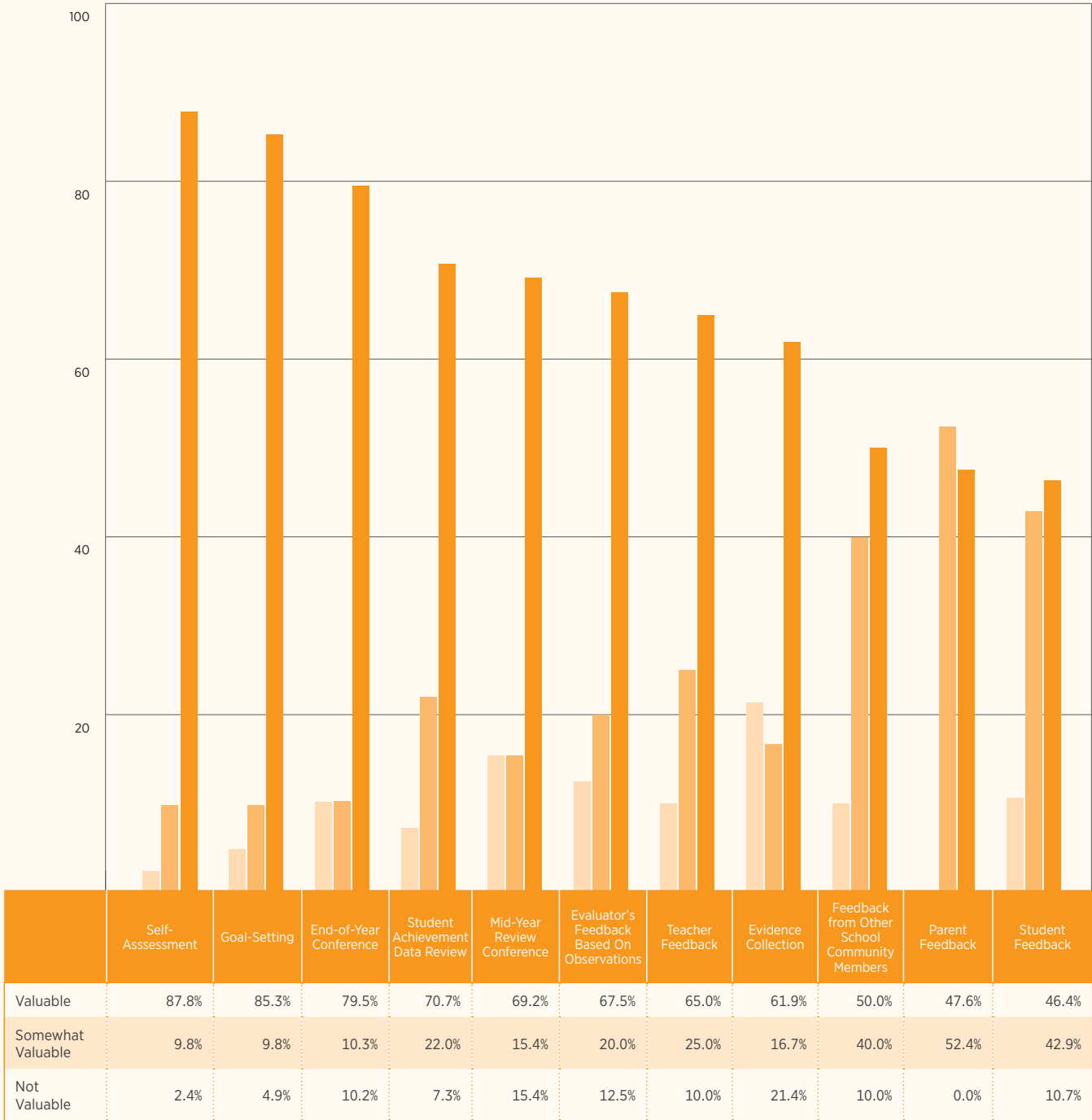
On the other hand, the stakeholder surveys received mixed reviews. Although 7 of 10 principals found evaluators' feedback from observations and interactions with members of the school community valuable to their professional growth, only 5 of 10 principals and 4 of 10 evaluators thought that parent and student feedback were valuable to principals' professional growth. However, it is important to note that only teacher stakeholder surveys are required; half of principals who participated in the end-of-year survey said they had no feedback from parents, students, or non-faculty members of the school community.

Principals and evaluators generally believed that evidence collection from a variety of available sources was valuable (for example, stakeholder feedback, documentation of professional development, student outcomes data). However, it is important to note that one fifth of the survey respondents (21.4%) did not perceive evidence collection as valuable. (See Figure 4.)

In the interviews, we probed for specifics on what evaluators and principals found valuable in the model. Evaluators noted that the conferences and reviews they held with principals in the course of the year allowed them to engage in analytical, reflective, and probing

FIGURE 4.
Perceived Value
of Various Elements
of the Example Model

SOURCE: Principals' Survey



conversations about leadership practice. Evaluators recognized the value of getting both principals and evaluators on the same page, understanding the professional strengths and weaknesses of each principal and developing a jointly agreed-upon plan of action to follow in the course of the year.

The fall conference was narrative and analytical, the mid-year probing and reflecting and then the summative was probing, reflecting and giving specific feedback on each of the components. This was more powerful than if we had stuck to a numerical checklist.

That [goal setting] helped me get an idea of where the principals are and where they want to go. Being part of that process helped us get on the same page and narrow down the focus of where the principals want to go and then develop a plan to help them achieve those goals.

Through these conversations with their evaluators and with the help of the performance measures and indicators, principals in the case study districts were able to develop a common understanding of their role and of the professional standards.

...the principals had not really had a conversation about what it means to be a principal. Repeatedly principals had told me, We finally have a license to be, to do the work that we are supposed to do. One of the most valuable parts of the process was getting a common understanding of what they should expect of themselves and what the district as a whole expects of them as well.

Similar to the evaluators, the case study principals most valued their engagement in the reflection process. Beginning with the self-assessment and goal-setting, these principals valued the opportunity to think through the available data and set specific goals as well as to track progress jointly with their evaluators.

The reflection pieces were most important—when I would sit down with my evaluator and set goals, and talk about how and what I would do to reach those goals. And then coming back mid-year and end-of-year and then looking at my goals and reflecting on how I did in reaching those goals and what I need to do to move forward.

Principals in the case study districts also highlighted the benefit of principal evaluation in aligning their annual goals with student achievement and school improvement goals.

In my previous district my evaluation wasn't tied to how the students were doing. Connecting the evaluation directly to what's going on in the classroom in terms of student achievement challenges the administrator to say what about my daily operation of the school as well as what are my long term goals and vision for the institution. How does that tie back to student achievement? Ultimately, that's why we're here.

Principals particularly valued that the feedback and ratings they received were based on observable data and not just a superintendent's opinion.

The big difference, and maybe that's the main difference between this model and the others that we've been evaluated under, was that the final evaluation was based on observable data versus opinion.

Principals also noted that their district's use of the model helped raise standards for their professional practice as well as provided them with focus areas that they could work on in the course of the entire year.

This is by far, for me, the most professional work that I've done for myself, ever being a principal. Part of it is the rubric that was created. The rubrics have high standards, and it made me have a target where I didn't have a target before. So that's been very helpful for me to see, "What am I shooting for here?" Instead of just thinking I'm doing a good job and have someone saying, "Yes, you're doing a good job," but have no meaning to it really.

It [the principal evaluation model] puts everything in perspective. It shows me the goals that I'm working towards. Obviously, I'm going to be taking care of all the aspects of being a school principal, but there are certain goals that we put out front that we were able to identify based on our stakeholder survey. Once we did that, then I've got the big things to focus on. It's a lot more detail oriented. I really like that aspect.

Finally, principals in some case study districts recognized that their district's model helped develop a professional community among principals and allowed them to adopt a more systemic and value-driven approach to the assessment of their practice.

This process has helped us to be more collegial, helped us to really look at our system and to see "what are the things that we value?" If I'm going to do this job what do I want people to see? What do I want them to judge me on? So I think that part, again, has been a valuable part of the process—looking at it, taking the time to really, line by line go through it to see what is it really saying, what is it asking us to do?

4. On which performance measures and other topics did the principals receive the most feedback?

The final survey asked evaluators and principals about the extent to which the feedback focused on seven areas:

1. *Instructional leadership*
2. *Vision and mission*
3. *Professional and ethical relationships*
4. *Human resources*
5. *Resource management*
6. *Adult learning and professional development*
7. *Community engagement*

a) Evaluators' Feedback to Principals

Survey Findings

Through the survey, evaluators reported that they most commonly offered feedback in the areas of *instructional leadership* and *vision and mission*. They were *least* likely to give feedback on adult learning and professional development, resource management, and community engagement (see Figure 5).

Case Study Findings

The case study interviews probed for specific examples of the kinds of feedback provided to principals. Similar to the survey respondents, the evaluators in the case study districts also said they emphasized feedback to principals on *instructional leadership*. Specifically, during the mid-year conferences, evaluators were looking at interim achievement results and were coaching principals on how to handle pockets of students who were not making gains in reading, how to coach teachers on better instructional strategies, and how to deepen implementation of adopted instructional programs.

Again, similar to what was shown by the survey data, evaluators in the case study districts supported principals in developing their school's *vision and mission* to align to those of the district. This was especially important in a district that had taken on a new vision and

¹³ The use of the word "novice" here refers to superintendents who served in a given district for two years or less. In other words, a superintendent who has worked in this capacity for a number of years can be considered novice in this analysis if he/she took a job in a new district within the past two years.

FIGURE 5.

Evaluators' Average Rating of the Extent to which they Provided Principals with Specific Feedback

(Note: the higher score denotes a greater deal of feedback provided in a given area)

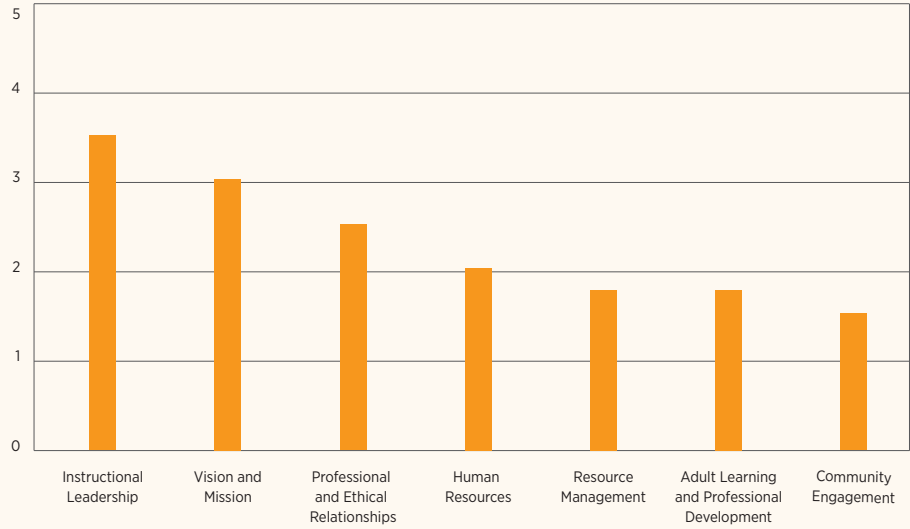
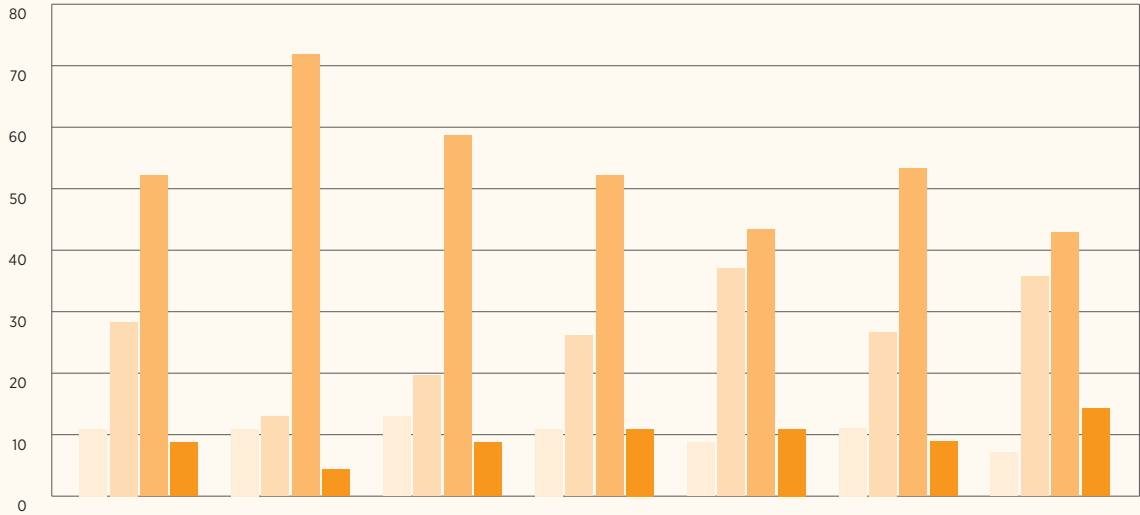


FIGURE 6.

Value of Evaluator Feedback Offered During End-of-Year Conferences

SOURCE: Principals' Survey



	Vision and Mission	Instructional Leadership	Professional and Ethical Relationships	Human Resources	Resource Management	Adult Learning and Professional Development	Community Engagement
Not Valuable	10.90%	10.90%	13.00%	10.90%	8.70%	11.10%	7.10%
Somewhat Valuable	28.30%	13.00%	19.60%	26.10%	37.00%	26.70%	35.70%
Valuable	52.10%	71.80%	58.70%	52.10%	43.40%	53.30%	42.90%
No Feedback	8.70%	4.30%	8.70%	10.90%	10.90%	8.90%	14.30%

mission as well as in a district that wanted to reinforce an existing successful vision and mission.

Some of the evaluators provided feedback on *professional and ethical relationships* such as coaching principals to deal with parents' complaints or to reflect on questionable decisions and helping to build morale in their schools by being transparent.

On professional and ethical leadership, my focus for our principals was that I wanted everything to be transparent. One of the things that came out of that is I wanted our high school principal to put out a newsletter once a week that went out only to staff and had a little tidbit on there that talked about how we can improve instruction once a week. The newsletter explained a lot of the upcoming events, what was going on. [Our goal] was to break down the barriers that were out there where they didn't trust administration and to make them feel we were on their side.

With regard to *human resources*, evaluators said they wanted to know which teachers were struggling—and how to help them as much as possible—while holding them accountable. They also wanted to know that their principals were building collegial relationships with teachers and breaking down barriers between the principal and teachers. Evaluators also raised issues related to hiring and discipline.

Evaluators reported having the least involvement with principals over *resource management*. Although most principals had individual budgets, usually they were small and intended to be used for principals' membership in professional associations and professional learning. In one case study district, however, principals played a strong role during budget reduction events by making suggestions of where to cut budgets.

b) Principals' Perceptions of their Evaluators' Feedback

Survey Findings

Principals' responses to the surveys support evaluators' assessment of the extent to which they offered feedback on performance measures. The vast majority of principals (84.8%) found feedback on instructional leadership received during the end-of-year conference at least somewhat valuable (see Figure 6). Overall, principals found mid-year and end-of-year feedback valuable to their professional growth. Nevertheless, 14.3% of the principals said they did not receive any feedback on community engagement; 10.9% said they had no feedback on human resources and resource management during the end-of-year conference.

Principals' overall perceptions of the value of feedback they received from their evaluators stayed stable throughout the year. On both the mid- and end-year surveys the principals thought that the most valuable feedback they received was in the areas of instructional leadership and professional and ethical relationships.

Case Study Findings

The case study interviews provided more insight about principals' sources of feedback on the performance measures. After the mid-year conference, principals reported that much of the feedback they received on the five Minnesota performance

measures emerged from the stakeholder surveys and from their initial goal-setting conference. Principals also received feedback on their professional goals and school improvement goals. To the extent a professional goal was related to one of the five performance measures, principals also received feedback on them.

Because establishing vision and mission is part of the principal's school improvement plan, the case study principals received feedback as they went through the process of developing the plan.

We have spent an extensive amount of time looking at mission and vision in our district, and the need to have a common purpose across all of our buildings. Through district in-services we've been given nominal group processing skills to help our schools arrive at a consensus around mission and vision.

All case study principals reported that they received feedback on instructional leadership through the review of their school improvement goals and feedback opportunities through district supports and monthly administrative meetings. As one principal explained:

We have an ongoing training every month. One has been on instruction to close the achievement gap—how to significantly impact our underserved learners. It isn't just the expectation that we develop action plans in our building but we are also getting educated as to the "how" and "why" for that.

Principals' reported that feedback on professional and ethical relationships was most often focused on building community within and outside the school and having integrity in relationships.

The principals often took human resources to mean hiring new staff, something that was rare because of several years of budget reductions. However, one principal saw it as a request from staff to spend more time in their classrooms.

From feedback on my survey I learned that the staff, even though they saw me more, they would really like to see me a lot more. They want me in their classrooms and they want me always around. This is something I want to improve on and can improve on."

Budget reductions and differences in allocation of resources to site-based budgets across the pilot districts limited most of the principals on resource management. But one principal who did have a budget described it this way:

I think the key is to acknowledge and recognize that I do have a budget, and I'm responsible for that budget, and whether it's the same as it was last year or more, or less. I have input into the cutting of the budget, but I don't need to own that. What I need to own is what's left over in the instructional climate, and the implementation process so that the implementation is consistent with our mission, vision, goals, and all that. My responsibility is to take the money you give me, and make sure our goals can get implemented.

5. How did feedback to principals vary by district characteristics?

The survey responses show that the perceived value of evaluators' feedback varied across different types of districts.

Examples of Variations by District Size and Student Achievement

Principals in small districts, on average, found feedback on adult learning and professional development, vision and mission, and human resources more valuable than did those in large districts. In addition, on average, principals in low-performing districts rated the feedback they received on vision and mission as more valuable than did principals of high- and average-performing districts. On the other hand, principals in high-performing districts considered feedback on community engagement as highly valuable (see Figure 7).

Examples of Variations by District Superintendent Tenure and Student SES Status

Overall, principals in districts with novice¹³ superintendents rated feedback they received during the end-of-year conference higher than principals in districts with longer-serving evaluators. The greatest differences in average ratings occurred with regard to feedback on adult learning and professional development, community engagement, and vision and mission.

However, although principals in high- and some-poverty districts had higher average feedback ratings across all areas, when asked how much they agree with the statement, "I received valuable feedback from my evaluator during the end-of-year conference," principals in low-poverty and rural districts were more likely to agree than principals in any other district types (see Figure 8).

¹³ The use of the word "novice" here refers to superintendents who served in a given district for two years or less. In other words, a superintendent who has worked in this capacity for a number of years can be considered novice in this analysis if he/she took a job in a new district within the past two years.

FIGURE 7. Perceived Value of the Feedback Received in Various Areas during End-of-Year Conference by District Size and Student Achievement

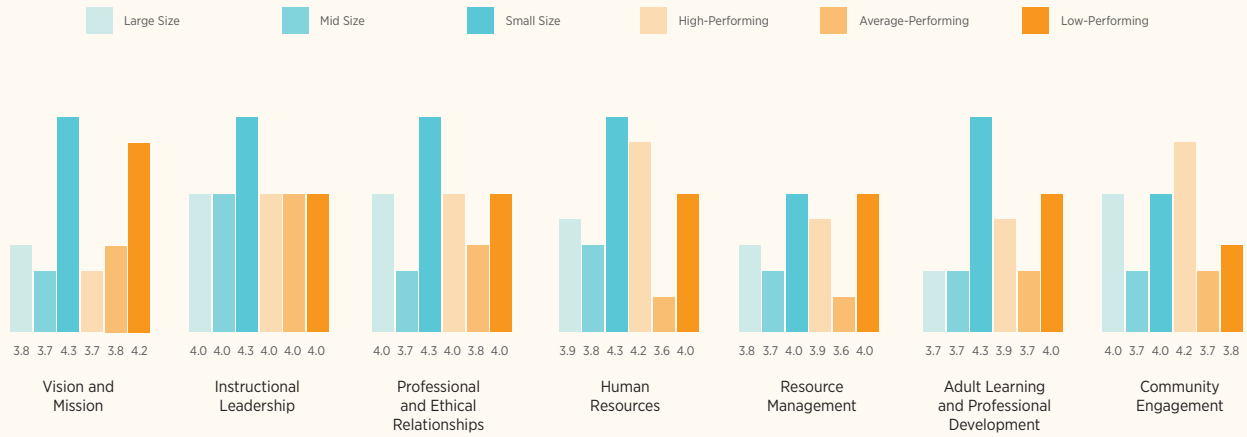


FIGURE 8. Perceived Value of the Feedback Received in Various Areas during End-of-Year Conference by Superintendent Tenure, District Urbanicity, and Student Socioeconomic Status



6. What is the value added of participation in principal evaluation for principal's professional practice?

In the end-of-year surveys, more than 80% of the principals reported that their participation in the pilot was most useful to them in the area of providing instructional leadership for high student academic performance. However, 31% said that participation in the model did not help them improve in establishing a vision and mission focused on shared goals and high expectations. Also, 26% said the process did not help them with building professional and ethical relationships through collaboration and effective communication (see Table 2).

TABLE 2. Principals' Assessment of How Helpful Participation in the Pilot Implementation of the Example Model Was to Their Improvement in Various Areas of Professional Practice

Areas of Professional Practice	Not Helpful	Somewhat Helpful	Helpful
Provide Instructional Leadership for High Student Academic Performance	19%	26.20%	54.70%
Building Professional and Ethical Relationships Through Collaboration and Effective Communication	26.10%	26.20%	47.70%
Establishing a Vision and Mission Focused on Shared Goals and High Expectations	31%	28.60%	40.50%
Managing Human Resources for Quality Instruction and Professional Growth	25%	37.50%	37.50%
Strategically Managing Resources for Systematic Performance Accountability	25%	40%	35%

Below we examine responses to questions that asked case study informants about the utility of the three main data sources required by the example model.

Stakeholder Feedback:

Several of the case study evaluators and principals emphasized the value of collecting stakeholder feedback. Some districts were able to collect stakeholder feedback in time for the mid-year conference and focused their discussion primarily around the findings from the stakeholder survey. One district utilized stakeholder (parents, students, and teachers) survey data from the prior year during the goal-setting phase of the principal evaluation process. Surveys that were carefully aligned to the Minnesota performance measures were perceived as providing the most value to principals.

Both principals and evaluators said in interviews that they were not clear about what survey tool they were expected to administer to collect feedback from various stakeholders. Many principals were dissatisfied with the stakeholder surveys eventually used in their districts, emphasizing that the questions did not provide valuable positive feedback about practices that should be continued and were often poorly understood by the school staff.

[I was] very disappointed that teachers were the only stakeholders that were allowed to give feedback, that I had no input into [choosing] the teacher survey, and that teachers were not given an opportunity on the survey to provide positive feedback (unbelievably, there were only spaces provided for negative feedback).

Achievement Data:

Both principals and evaluators criticized the late availability of the state MCA and MMR data. Interviewees said that the timing of the student achievement data release in late September delays principals' ability to assess progress toward their school improvement goals. In addition, the need to wait for these data extended the principals' evaluation cycle past June 30th and into the new contract year.

Portfolio Evidence:

Finally, a small number of the case study principals said their evaluator did not give them an opportunity to discuss their performance evidence (often portfolios of materials housed in notebooks) they had collected in the course of the year and, therefore, did not engage in deep reflection about their professional practice and school performance.

I've revamped that for myself [portfolio evidence] and talked through it with [my evaluator] — that it felt weird and there was not an accountable piece for me. It was, "Oh, I put all this work into it and you don't even want to look at it." So then I thought okay, if I'm going to continue to gather documentation, how is it going to be useful for me and not putting a binder together was the answer. So I've scaled that back and I'm doing things just electronically. I'm still finding myself asking questions. Is this evidence of what I'm doing?

7. What was the perception of the capacity of various data sources to measure principal performance accurately and fairly?

Overall, the survey responses show strong alignment between the perceptions of evaluators and principals about the most accurate sources of data, including evidence of progress and outcomes of their professional growth goals, direct observation of principals' practice, self-assessment results, and mid-year conference results and feedback. The one variance was in the perception of direct observation as fair; this discrepancy in the perceived fairness of direct observations between evaluators and principals needs further study (see Table 3).

Evaluator Case Study Findings

The case study interviews provided a better understanding of evaluators' perceptions of the accuracy and fairness of different data sources. Evaluators made the following six common observations about the data sources.

District Modifications to the Model increased the Perception of its Fairness and Accuracy.

Evaluators thought the model was fair and measured principals' performance. However, they believed that sometimes the modifications and processes districts created enhanced the model's fairness. One district evaluator chose to use a principal self-evaluation process based on a narrative that was scored holistically. In the goal-setting process, this provided the basis for a dialogue between the principal and evaluator, leading to selecting appropriate goals, measures, and action plans. It also allowed the evaluator to more closely track each principal's progress. Two other districts modified the rubrics and the descriptors for each indicator, clarifying in which performance level principals should be placed.

Change in the MCA Reading Test Undermined the Fairness of the MCA Reading Goal.

All evaluators reported what they believed to be a significant mitigating factor to the fairness of the model. This was the change in the state assessment program, the Minnesota Competency Assessment (MCA), which recently had been aligned to the more rigorous standards of the Common Core.¹⁴ Pilot districts reported their literacy proficiency scores were about 15% below the previous year's scores, making both fair cohort comparisons and growth scores impossible to determine.

¹⁴ The Common Core State Standards are designed to be robust and relevant to the real world, reflecting the knowledge and skills that our young people need for success in college and careers. These have been sponsored by the National Governors' Association and the Council of Chief State School Officers and have been adopted by 46 states.

TABLE 3. Principals and Evaluators' Perception of How Accurately Each Source of Evidence Used During the Entire Evaluation Process Assesses Principals' Performance

Source of Evidence	Principals' Mean Accuracy Score (of 5)	Somewhat Helpful
My Principal Self-Assessment Results	4.08	92.3%
Evidence of Progress and Outcomes of my Professional Growth Goals	3.95	100%
Mid-Year Review Results and Feedback	3.88	92.3%
Student Growth Data	3.83	84.6%
Student Achievement Data	3.78	61.5%
Documentation of the Professional Development Completed During the Year	3.74	76.9%
Direct Observations of my Practice	3.71	100%
Indirect Observations of my Practice	3.65	84.6%
Teacher Feedback	3.58	69.2%
Feedback From Non-Teaching School Staff	3.35	7.7%
Evaluator's Notes from Interactions with Teachers	3.32	46.2%
Parent Feedback	3.30	23.1%
Evaluator's Notes from Interactions with Other School Community Members	3.21	15.4%
Evaluator's Notes from Interactions with District Staff	3.15	23.1%
Student Feedback	3.12	23.1%

Note: The evidence types are arranged from most to least accurate in their capacity to assess principals' performance as perceived by principals. The mean accuracy score ranges from 1 "least accurate" to 5 "most accurate". In other words, the higher the score, the most accurately principal believe a particular type of evidence captures their performance.

Late Reporting of the MCA Data also Complicated and Delayed the Evaluation Process.

Although the use of MCA data is not required for the 35% student achievement measure, the interviews revealed that districts wanted to use these data to help set goals for the coming year and determine related professional learning opportunities over the summer. Both principals and evaluators expressed disappointment over the late timing of the MCA results which are not reported to schools until late September. As a work-around, one district used interim assessment results as proxies for school improvement measures to finalize principals' evaluation ratings in June—then either principals were given the full number of points for reading achievement in the rating process or the goal was deemed not applicable because the 2011–2012 MCA reading scores could not be compared to those for the previous year.

I. Rubrics Determining Performance Levels Were Unwieldy.

Another issue raised in multiple interviews was the clarity in the rubric descriptors that assessed principals' skills on the performance measures and rubrics. Both evaluators and principals reported it wasn't clear where they "landed" on a particular rubric and that this was the most unwieldy part of the process. Principals agreed that the dialogue around the rubric was the most valuable, and coming up with a quantitative rating was least valuable. One principal reported that the rating score didn't help them know what to do to improve. Two pilot districts developed solutions for this problem, one by rewriting the rubric descriptors and another by working with principals to reduce the 26 indicators to 10 "power standards" with more explicit rubrics. The former district convened a subgroup of principals and administrators to revise the rubric descriptors within each of the performance levels to be both more explicit and aligned to the school improvement strategies being used in that district.

ii. Setting Consistently Rigorous Goals Sometimes was Challenging.

Another threat to fairness of the evaluation model was that some principals set more rigorous goals for themselves than others. For example one principal might set a goal of increasing reading proficiency by 2% and another by 10%.

As a principal I can set goals two ways. I can set goals that stretch my organization and stretch learning, or I can set goals that will make me look good in the principal review process so I keep a job. And I will never set goals for job security. I'll set goals for my building because that's what I'm all about.

Confounding this problem is a school's current level of proficiency. If, for example, the proficiency level for reading is set at 40% then an 8% gain might be possible. But it is much more difficult for a school with reading proficiency at 95% proficiency to increase even a couple of percentage points. One district solved this by trying to ensure each school would reach proficiency by 2020¹⁵ and setting annual gains for each school and subgroup to achieve this goal. Another district plans to create a panel of district staff to help calibrate goals across principals and their schools through dialogue with them. When principals present their goals and targets to the panel, the district can also identify the supports principals will need to reach their goals.

¹⁵ Minnesota reset the No Child Left Behind (NCLB) goal of reaching proficiency from 2014 to 2020.

iii. Knowing the Principals Well Takes Time

Finally, two of the evaluators were relatively new to their positions and didn't feel they knew the principals as well as they would have liked. This led to some unease on their part as to where to place principals on the rubric during the initial self-assessment conference. One evaluator reported resolving this by mostly agreeing to the principals' initial self-assessments and spending more time observing them in their schools during the school year.

a) Principals' Mid- and End-of-the-Year Responses

Principals' survey responses to the questions about accuracy and fairness of different data sources were by and large consistent across both periods of data collection. Following their mid-year conferences, principals selected self-assessments, evaluators' observations and feedback, and artifacts or principals' portfolios as sources of evidence with the highest capacity to accurately and fairly capture their performance. Respondents gave the lowest ratings to focus groups with stakeholders as able to assess principals' performance in an accurate and unbiased fashion (see Table 4).

Principals' responses to the end-of-year survey indicated that they continued to view their self-assessment results as able to provide the most accurate assessment of their performance; student feedback was seen as the least accurate assessment. The principals also gave favorable responses to the accuracy of evidence of progress, outcomes of principals' professional growth goals, and mid-year results and feedback (see again Table 4).

Table 4. Assessment of the Capacity of Various Types of Evidence with Regard to their Capacity to Accurately Capture their Performance

Evidence Considered During the Mid-Year Review	Mean "Accurate and Fair" Score (Of 5)	Number of Valid ¹⁶ Responses
Principal Self-Assessment	4.38	42
Evaluators Observations and Feedback	4.26	35
Artifacts or My Portfolio	3.91	34
Student Growth Data	3.83	40
Student Achievement Data	3.78	40
Parent Survey	3.44	23
Teacher Survey	3.43	30
Student Survey	3.26	19
Community Partner Survey	3.13	15
Focus Group with Community Partners	2.57	7
Focus Group with Students	2.57	7

Below we highlight principal responses during the case study interviews concerning four key kinds of evidence.

i. Self-assessments

As the survey results show, the vast majority of principals who completed the end-of-year conference thought their self-assessments provided the most accurate evidence of their performance. Case study interviews illuminated the thinking behind these results.

¹⁶ Valid responses were those other than NA.

I think that the chance to self-assess was important for us to reflect on where we currently see ourselves and what are our strengths and what are our opportunities for improvement. It forced us to take that time to do it because, as you know, when you're running, that reflection time really can get pinched. And it's something you can put to the side because other things that you may think are more important rise to the top. So it forced that to happen.

li. Stakeholder Survey Results

In the interviews, principals had mixed responses to the stakeholder surveys. Some principals put significant emphasis on their results, but other principals were concerned about the survey administration, survey format, and the alignment of the survey questions to the performance measures. Many principals had concerns about the survey data from teachers and parents (where parent data were collected). Some worried that results could depend on a recent event in the school, the timing of when the survey was administered, and/or the percent of those who responded to the survey.

Where there were high response rates, such as 80% to 90%, principals put credence in the survey results. One case study district piloted the survey in just one school early in the school year after the school had gone through some upheaval during the prior school year. This gave district leaders and the principal timely information on teachers' perceptions of school leadership and school climate. The principal used the results to make immediate changes in the school and then used them as a baseline for the following year.

In a different example, it appeared that PLCs had become so ingrained in the school culture that teachers didn't see them as professional development and marked their opportunity for professional development low on the survey. Stakeholder survey results were better received if the district had previously administered the survey, allowing for baseline data, or had administered a stakeholder survey with national or state norms for comparison purposes. One customer-driven district used their prior year stakeholder survey results to set the survey target satisfaction rate for all schools at 90%.

Another principal was disappointed when he didn't meet one of his professional goals because the survey response format had changed. He set his goal based on a prior survey, which had five Likert-type responses from "strongly agree" to "strongly disagree." When the current survey was administered, the central office had reduced the number of possible responses to just three, and only the number of responses in the first category was counted toward meeting his goal.

The high school principal job I think is very complex, and it's very important that the proper weighting is given to the proper categories, so it comes out and that complexity is represented. An example for me would be one of my goals was to improve PLC performance and the perception of teachers, as far as that being meaningful staff development. We've made progress, but I had set a high goal, in the 90s, and we were probably in the low 70s or high 60s as far as teacher perception. I didn't meet my goal, so I got a zero on the ranking. I didn't get credit for partially meeting it.

iii. Late Arrival of the MCA Data

As discussed earlier, nearly all case study principals shared concerns about the late delivery of the state assessment results. Because the evaluation model recommended the use of the MCA data (which didn't arrive until late September), the majority of districts delayed the final end-of-year conference to the fall. (Nonetheless, several districts conducted a partial end-of-year conference in June or July and then finalized the review in October after the MCA results became available.)

The four case study districts dealt with the late MCA data in various ways. One district closed the books on the evaluation in June, using some of their interim assessment data to gauge student proficiency and growth. Two districts rated two components of the model—stakeholder survey results and the evaluators rating in June—then waited until October to evaluate the final component, student achievement. One district waited until October to undertake the principal's rating in all three evaluation components.

iv. Rubrics Determining Performance Levels

Many principals and evaluators reported that the rubrics were sometimes dense and difficult to use. Two districts revised the rubrics to add more clarity and fairness to the evaluation. Principals had some confusion about what the rubrics meant, especially moving across the performance levels. Once the rubrics were revised to add more clarity to each of the performance levels and were aligned to specific district expectations and principal practices, principals had more faith in them.

We took the state information and we created some power standards based on a system that we were using, so we ended up with about ten what we're calling "power standards" and aligned them with this system. They better captured what we were trying to get at here.

V. Other Thoughts on Accuracy and Fairness of the Principal Evaluation Process

Principals in the case study districts also thought the evaluation system was fair when they had some control over the outcomes. Goal attainment either had to be supported by the district or under the principal's control. It also helped when goals were realistic and attainable.

My rating was fair based on the data that we had in front of us. Unfair, because there were things that were out of my control. So I accepted it and decided that I'm okay with it as long as it's baseline data and that we've already put systems in place that I'm believing will change the outcome of that.

Principals tended to have more confidence in the evaluation process when their own ratings of their performance matched the evaluator's rating.

I think the evidence I have is the evidence I collected about my own self-evaluation and it matched with what my supervisor was saying as well. So while he didn't have the evidence that I had pulled out [a portfolio] which included meetings and documentation, that kind of thing, his assessment of me using the rubrics was similar, if not completely the same. There wasn't a place where there was a disagreement really. In fact, he always put me higher than I thought I was on some of the areas.

8. What skills do evaluators need in order to provide effective feedback to principals?

In the first round of case studies interviews, we noticed that districts that seemed to have the most effective implementation of the model were those where evaluators and principals described a “non-threatening environment” for conversations between evaluators and principals. This led us to want to probe further on what makes evaluators effective and able to build trust to promote professional growth. Consequently, a question was added to the second round of case study interviews of both evaluators and principals to collect their insights.

We found broad agreement between evaluators and principals in the case study responses on effective evaluators:

They had deep knowledge and understanding of a principal’s job and skills.

They knew best educational practices and how to give nuanced feedback and support to principals—even experienced principals. They were knowledgeable about district programs and how these could support principals in their schools. They also knew what was going on in each school. “You need to be fluent in what’s going on in each school on a daily basis. That only takes place by having regular, personal connections with not only the leader of that building, but people within the building. I think that’s how you gain knowledge.”

They had excellent communication skills, including being able to have difficult conversations in a way that provided support to principals, which often required having courage and compassion.

Most interviewees agreed that good evaluators were good listeners, and were able to reflect and discuss principals' performances with them and ask good questions to help principals grow.¹⁷

A principal: "You should be able to give somebody positive and negative feedback and in each circumstance and the person should be okay with that. When the evaluator leaves the room, you should not feel beaten down."

An evaluator: "I think it's to know when that person needs to be called out—and as hard as that is—and we did that—to say, 'This isn't going to cut it. Right now we're okay, but this is not going to cut it long-term in these particular pieces.' So delivering those difficult messages with grace, I think, and an understanding of support. And that we're not here right now to talk about anything other than we're here to support you. I also think the other skill is to know when to use a different tool in order to start down another road of employment consequences. And thankfully, we're not at that point, but I see that being an important skill for this."

They knew how to build trust and rapport with principals.

One principal said, "He has my back and I know we are a team." Effective evaluators did what they promised and showed integrity and honesty with their principals. Some worked tirelessly to help principals understand what the performance measures looked like in practice so the summative evaluation would have full transparency for the principals.

If a superintendent doesn't have a rapport with the building principals, if the superintendent doesn't have the credibility and integrity with the principals, it's going to affect the evaluated person. It's going to be hard to earn the trust so that the suggestions and the improvements are taken the right way.

They were organized, data-driven, and results oriented.

Effective evaluators streamlined the principal evaluation process, always used data and evidence in their coaching conversations, and kept principals focused on their goals. "Evaluators need consistency, efficiency to minimize the managerial parts of their jobs so they can have the time to do observations to inform the system. Are you effective enough to make sure you're following through and doing everything you've promised to do?"

They were reflective, constructive, analytic, thoughtful, and mindful.

They were able to construct conversations with principals that helped them understand what it meant for their performance as a school leader aligned to the performance measures.

¹⁷ Several districts had evaluators and principals trained in a model called "cognitive coaching." For more information see Arthur L. Costa and Robert J. Garmston, *Cognitive Coaching: A Foundation for Renaissance Schools*, Christopher-Gordon Publishers, Inc., 2002.

And being able to accurately reflect in writing and in conversation for people what do those things mean for your performance as an administrator? So being able to take that data, both observed and human and those kinds of pieces with the hard test score data, and being able to articulately put that in a way, in written form, that aligns with the expectations of the rubrics. And being able to explain that in a way that is differentiated for each person.

One principal summed it up well:

They have the ability to ask the right question, an ability to direct the conversation, an ability to go with key points that you need out of that meeting and to get the principal to understand. Discussion and coaching are how you create change.

Summary Comments

In summary, evaluators generally agreed that the model can measure principals' performance effectively if thoughtful measures are used and the final rating calculation is fair. In particular, evaluators noted that the self-reflection on the part of principals was a major factor in effectively measuring a principal's performance.

Although principals generally found implementation of the example model valuable to their professional improvement, they emphasized that the following three conditions need to be in place in order for the model to produce the best outcomes for their professional growth and school improvement:

- » *The example model is understood and implemented as a professional growth model, not a simple accountability system.*
- » *Evaluators are trained to effectively implement the process.*
- » *Principals and evaluators implement the model as a collaborative team.*

These and other recommendations from principals and evaluators will be discussed in detail in Parts III and IV of this report.

PART III.

**IMPLEMENTATION OF
THE EXAMPLE MODEL
AT THE SCHOOL AND
DISTRICT LEVELS**

GIVEN THAT THE model designed by the state is intended to be an example of the principal evaluation process, some of the pilot districts chose to adapt the model to better address their specific needs. Half of the evaluators reported on the survey that their districts were either making significant adaptations to the state’s example model or using their previous evaluation model. Those superintendents who chose not to adopt the example model believed that their district’s principal evaluation models met or exceeded the state’s model in terms of its effectiveness in supporting principals’ professional growth and accountability for student results.

The following four study questions examine implementation, including variations, challenges and support conditions.

1. How well did principals and evaluators understand the example model?
2. What did the implementation of the example model look like at the district level and how did it vary?
3. What challenges were faced in implementing the principal evaluation model?
4. What conditions support successful implementation of principal evaluation?

1. How well did principals and evaluators understand the example model?

To ensure familiarity with the example model, in the beginning of the pilot year both principals and superintendents participated in orientation training provided by MDE. The mid-year survey asked them to reflect on how well the orientation sessions prepared them to implement the model.

On average, principals and superintendents alike reported on the survey that they received sufficient training and felt prepared to implement the example model with fidelity. The area where both principals and evaluators felt least prepared was evidence collection, a critical element of the example model if it is to be effective for mid-year formative and end-of-year summative conferences. Evaluators also did not feel adequately prepared to conduct principal observations, which raises important questions about evaluator preparation and supports.

Principals in low-poverty and high-performing districts reported higher levels of preparedness to implement the example model. Principals in these districts also were less likely to express concerns about their evaluator's capacity to coach principals through the self-assessment and goal-setting stages of the process.

2. What did the implementation of the example model look like at the district level and how did it vary?

Most districts engaged their principals in self-assessment and professional and school improvement goal setting in the fall and conducted mid-year formative and end-of-year summative conferences. According to the end-of-year principals' survey, the most commonly used elements of the example model included orientation, self-assessment and goal setting, feedback from evaluator, the mid-year formative conference, and the end-of-year conference. The case study interviews, summarized below, focus on the nature, quality, challenges, and variations in the implementation of these elements of the example model.

a) Self-Assessment and Goal-Setting

Commonly, principals in the case study districts engaged in the self-assessment and goal-setting process sometime between late August 2012 and early October 2012, following their participation in the August state orientation trainings. As the first step in the process, evaluators held administrative meetings with their principals to further define the expectations for self-assessment and goal setting, to review the state-provided documents and recording forms, and to clarify how to undertake the self-assessment and how to set school improvement and professional goals.

Principals' goal-setting meetings were typically scheduled a few weeks after the administrative meetings. Prior to these conferences, most principals shared their self-evaluations and

**a) Self-Assessment
Goal-Setting Cont.**

draft goals with their evaluators. The conferences, which took place as a dialogue between the principal and the evaluator, lasted about an hour and a half. In the interviews, the evaluators reported that they encouraged principals to develop adequate measures to monitor progress on their goals. This is how one principal described the process:

... they then allowed us about a week to develop our goal-setting process for both the self-assessment for growth and the school performance measures. Then we met again to review those, almost a draft situation if you will, where both of them read it from the perspective of are they SMART goals, are they measurable, are they representative of what the state is looking for in terms of the data that will reflect growth for our school and also as an administrator? They were also asking for clarification.

One of the evaluators interviewed had a similar recollection: “They were all measurable, all based on data, were achievable and they had very good supporting action plans.”

To identify principals’ professional growth goals, principals and evaluators mostly used student achievement data, district goals and progress measures, and self-assessment results. Districts that had a prior year’s stakeholder survey results available also used them to set professional goals. Often principals in the case study districts were asked to develop professional goals in their self-assessment around weaknesses or those areas identified as weak through the stakeholder surveys. Evaluators determined the number of goals and the categories for principals, usually in advance.

Because in most case study districts the upcoming year’s school improvement plans were developed before the principals’ goal-setting meetings, these plans influenced a principal’s school improvement goals. To the extent these were aligned, it seemed to streamline the evaluation process and increase the potential outcomes for district, schools, and principal.

To set school improvement goals, principals in the case study districts relied most heavily on student achievement data and district goals, especially student growth data. Both principals and evaluators thought that student achievement data, district goals, and self-assessment results were most useful in helping principals set their individual professional and school improvement goals.

Because this was a “low stakes” pilot implementation of principal evaluation, evaluators in the case study districts reported some flexibility in the goal-setting process. In several districts, evaluators reported either completely accepting the goals proposed by the principals or accepting them with just a few minor revisions.

In preparation for this first phase of the principal evaluation process, some of the case study districts provided principals with additional training (that is, training on SMART goals)¹⁸ while other districts adapted the self-assessment and goal-setting processes proposed by the state. For example, one case study district chose to use a narrative format for the principals’ self-assessment with a holistically determined score for each indicator. Another district discouraged principals from using their school building’s school improvement goals, but instead asked principals to articulate their role in achieving their school improvement goals

¹⁸ A SMART goal is Specific, Measurable, Attainable, Realistic, and Timely.

and measuring this. In one district the assessment director helped principals set their school improvement goals and the evaluator/superintendent helped them set their professional growth goals; other districts helped principals set realistic student achievement growth targets based on where they were and where they needed to be by 2020 to meet the NCLB goals.

Most principals and evaluators who responded to the survey found the self-assessment and goal setting phases of the principal evaluation process useful to the principals' professional practice and development. Nonetheless, principals' responses pointed to several areas that need further attention to improve the outcomes of this important stage of the principal evaluation process. For example, over one-fourth of principals disagreed that the recording form was useful in guiding the development of school improvement goals. Also 17% of principals disagreed that they had a clear action plan to follow as a result of the self-assessment and professional goal-setting phase. (None of the evaluators found either area problematic.) Principals in high-poverty and low-performing districts also were more likely to express concerns about the amount of support available to them during this stage of the process.

b) Preparation for the Mid-Year Conference

Following the fall principal self-assessment and goal-setting conference, the majority of principals who responded to the survey engaged in one or more of these activities:

1. Implementing the action plan for professional growth and school improvement.
2. Discussing expectations for the mid-year conference with the evaluator.
3. Receiving feedback from the evaluator on principals' individual professional growth and school improvement progress as well as implementation of the action plan.
4. Being observed and visited by the evaluator.

By the time of the mid-year conference, we found both through the surveys and the interviews that principals and evaluators were making significant changes in their practice and leadership development strategies.

On average, principals engaged in about half of the activities listed in Table 5 below. The number of preparatory mid-year conference activities fluctuated, however, based on the districts' poverty levels, size, and superintendent tenure. Smaller and high-poverty districts as well as those with a recently arrived superintendent tended to engage in fewer preparatory activities.

Table 5. Principals Engaged in Preparation for their Mid-Year Review

Preparatory Activities for the Mid-Year Conference	Percentage Of Respondents Who Participated in This Activity
The Evaluator Provided Feedback on The Evidence I had Collected	19%
The Evaluator Helped Me Plan for the Mid-Year Supervisor Review	19%
I Created a Portfolio of Artifacts to Document my Individual Professional Growth	20%
I Collected Interim Assessment Data on School Improvement Goals in Preparation for the Mid-Year Supervisor Conference	22%
I Made Revisions to my Individual Professional Growth Goals, School Improvement Goals, Action Plan(s)	24%
The Evaluator Conducted School Visits and Observations	29%
The Evaluator Provided Feedback on the Implementation of my Action Plan	30%
The Evaluator Provided Feedback on my School Improvement Progress	31%
The Evaluator Provided Feedback on my Individual Professional Growth	32%
The Evaluator Communicated the Expectations for the Mid-Year Supervisor Conference to me	33%
I Implemented the Action Plan(S) for Individual Professional Growth and School Improvement Goals, Collecting and Organizing Proper Evidence	39%

How the Case Study Principals Prepared for the Mid-Year Conference

Responses in the case study interviews with principals on what happened between the goal- setting prepared for the mid-year conference and the mid-year conference corroborate the survey results. First, the principals set up a strategy for evidence and artifact collection related to all of their goals. Second, they began to implement their professional growth and school improvement strategies.

We had a setup where it was going to be a portfolio, and when I did something I would put that into my portfolio and then we would review that with my evaluator. So I was putting staff agendas, any meeting agendas I had as far as our data review meetings, any plans that we had to make to come up with for changing instruction, data, monthly data that we had was put in there. I put in communication to parents; if I had to communicate with parents, letters I sent home or a narrative of conversation I put in there. I put our literacy plan in the portfolio. Schedules, I put schedules in the portfolio. Pretty much anything I had hard copies of, I put in the portfolio. Basically in mine I put almost everything I did in the portfolio and categorized it to show what indicators those fell under.

Most case study principals said they already had been helping teachers with securing and analyzing interim achievement measures of students’ academic progress. This became even more important because student achievement progress was also a school improvement goal of the principals’ evaluations. In two districts, principals dug even more deeply into the data to help teachers identify students or cohorts that needed more help in specific learning areas. Interim student achievement data were used to track student progress

on their academic goals. One district was able to use interim measures to predict student success on the state assessment, which created faster and deeper interventions for low-performing students during the academic year.

I picked this specific group of my fourth and fifth graders because their scores will have a profound effect on our overall testing scores. And then worked with [the evaluators] to really help define that and figure out, “Okay, which pieces of assessments do I need to have as artifacts to show that that is actually happening?” And how many [students] were far below, and how many were below and how many were at grade level? So I knew that—15 out of my 85 fifth grade students were far below. So if I take the 15 lowest ones, if we could have a positive effect on their test scores, I know that’s going to raise all of them up. So that’s why I collected those two artifacts to begin with. And then I knew which students I was going to be tracking. It was much more purposeful as well because knowing that these are my lowest students that need the most assistance then I knew that it could affect everybody. And I also knew that I had the tracking mechanisms.

Principals in the case study districts also reported that, as a result of their participation in the model pilot implementation, they were able to spend more time in classrooms and PLCs, ensuring that the latter was an effective mechanism for a school to meet its school improvement goals. Principals were paying much more attention to the means and strategies to achieve student achievement gains, even putting in place some accountability measures.

We spent two hours with our staff, again, every Wednesday, an hour each time where the collaborative teams, again, grade levels, needed to look at their data and their action plan in a formative way to see where they’re at. We’d see if they need to tweak their [PLC] action plan at all at this point, and then I needed to gather that to present at the mid-year conference.

One case study district had received its stakeholder data during this period, and principals used these data to make sure progress toward professional growth and school improvement goals was on track.

When we got the survey, we had further proof from our stakeholders of what we thought were our goals. It ended up pretty much right in line with what we came up with.

All the interviewed principals reported frequent interactions with their evaluators. Historically, most principals said they had frequent interactions with their evaluator based on what was going on in their school. Evaluators (usually the supervisor) would help principals to problem solve an issue or simply check in with them periodically.

We [principal and evaluator] talk all the time but not necessarily about evaluation. I would say the biggest thing we’ve had is knowing I’ve got support there. At this time, from between the goal-setting conference and the mid-year conference, I would say that the things that I’m working on are more building specific—working with the student support team. I meet with them every single week, and also I’m involved in the teachers’ professional learning communities as well. So I’m finding out, “Okay, what’s working?” in those areas, but I’m also finding out which students aren’t making gains and what we’re doing—what are we doing with those specific students?

How the Case Study Evaluators Prepared for the Mid-Year Conference

Consistent with principals' responses, the case study evaluators said they prepared for the mid-year conference by working with the principals to monitor and adjust their individual professional growth goals, school improvement goals, or action plans; communicating the expectations for the mid-year conference to the principals; and conducting school visits and observations.

Through interviews with evaluators, we also learned that they took a more active and systematic interest in visiting schools and speaking with principals. Evaluators often described their prior years' interactions with principals as "anecdotal." Under the new evaluation process, one evaluator made a commitment to visit every school and every classroom in the district during the academic year. Another committed to visiting every school and the principal several times during each trimester. Evaluators would typically provide either written or verbal feedback to principals on what they observed in their schools as explained by these two evaluators:

I meet with them [principals] once a week. I'm in their buildings. I'm talking to them about projects or programs or whatever, so I'm "dip-sticking" on a fairly regular basis, how is each building doing, how is each building principal doing.

I guess knowing that they're working their plans is what's most important to me. If they created their plans and their goals in the fall and then they don't think about them or revisit them until they come to my meeting in January, I would probably be very concerned. So we often do check-ins at our meetings around their school improvement goals and plans. I also have [another district leader] doing some kind of incremental check-ins with principals as well as my Teaching and Learning Department regularly to be sure that they're working their plans and implementing their plans and getting the support they need to implement their plans..

Evaluators in the case study districts supported district efforts to ensure principals had effective learning programs in their schools and interim student achievement data. Principals often collected interim data, and some showed it to the evaluator as it became available, allowing them to collaborate on solving issues as they arose.

Evaluators also reported informally monitoring principals' progress on their goals. Although principals' goals rarely changed during the year, the strategies to achieve them could be modified.

There is good data for me, as an evaluator, to say that this person is really thinking about their practice, they're willing to take risks, they know that sometimes those risks don't pan out, but they cannot only tell me why it's not panning out but they can tell me what is the next plan of attack, in terms of how they're going to do that.

Evaluators also let principals in the case study districts know when and how the mid-year conference would be conducted. Some evaluators looked at the mid-year questions provided by MDE and then created a shorter list of questions for the principals that they sent out in advance through an e-mail. Principals were alerted as to what evidence they should bring to the mid-year conference and whether to send it to the evaluator in advance.

I sent them the agenda so they have that and they have the questions I'm going ask. They know they need to bring me evidence around each one of those questions, and they end up almost leading it because they know what my expectations are.

These evaluators frequently tailored monthly or biweekly meetings with principals to school improvement issues or to helping the principals deepen their understanding of the evaluation process. One district sent out a monthly newsletter to the principals that documented promising practices seen during observations in the district's schools. Another tailored the performance measures and rubric to the district. In these two districts the evaluators were relatively new, but it was clear the evaluation model was driving conversations that deepened principals' understanding of the performance measures and the sharing of promising practices.

I do a monthly administrative newsletter and for all of them I summarized the best practices that I've seen and the ineffective things that I've seen in our district so that they could reflect on them. I've provided them with the same information in that regard.

Few of the case study districts had been able to collect parent and student feedback in preparation for the mid-year conference. As discussed in Part I of this report, this may have resulted from the lack of clarity surrounding which stakeholder surveys districts would eventually select and administer. However, most principals collected student assessment results, interim progress on meeting school goals, student growth data, and evidence of their professional development completed in the course of the year.

Over 70% of the evaluators said they also provided principals with the evaluator observations and feedback, although only half the principal respondents to the survey indicated having received this kind of feedback (Table 6). This discrepancy warrants further study.

Almost a third of all the principal respondents (32%) collected three to four types of evidence; another two-fifths of the respondents collected five or more types of evidence (Table 6).

TABLE 6. Types of Evidence Principals Collected and Evaluators made Available to Principals in Preparation for the Mid-year Conference

Types of Evidence Collected in Preparation for the Mid-Year Conference	Percent of Principals who Collected this Evidence	Percent of Evaluators who Report that Principals Engaged in this Activity (N=14)
Student Assessment Results	61.1%	63.6%
Interim Progress on Meeting School Improvement Goals	59.3%	27.3%
Student Growth Data	55.6%	63.6%
Evidence of the Individual Professional Development Complete	53.7%	27.0%
Teacher Feedback	46.3%	54.5%
Artifacts or a Portfolio	33.3%	18.2%
Evaluator Observations and Feedback	31.5%	72.7%
Student Feedback	16.7%	9.1%
Parent Feedback	14.8%	0.0%

The case study principals reported they were collecting and using interim student achievement measures to assess progress toward their goals. There was some confusion at the beginning of the pilot because the pilot districts were instructed to use only the MCA data for student proficiency and growth measures. This recommendation was relaxed during the later state-level orientation meetings, so pilot districts began to use their interim assessments to monitor progress on the principals' school improvement goals.

Mid-Year Review—Variations Between Principals and Evaluators

The vast majority of principals (84%) reported on the survey that they had their mid-year reviews during the months of January through March 2013. In their mid-year conferences, principals and evaluators mainly focused on the discussion of three types of evidence they collected: student growth data, evidence of interim progress on the school improvement goals, and documentation of professional development completed. Evaluators also said they gave feedback to principals based on their observations of principals' practice. Some anomalies occurred, however. This claim contradicts principal survey results, where only half of principals reported evaluators' feedback was a type of evidence discussed during the mid-year conference.

Though not a large share of respondents, 6% of principals said the review never took place or was informal at best. Some principals indicated on the survey that they were not aware whether they had a mid-year conference, which suggests that some evaluators might have followed rather loose or informal procedures, thus potentially depriving principals of an opportunity to engage in evidence-based, feedback-driven, and outcomes-oriented conversations around their professional practice and school improvement.

Nearly three-fourths (74%) of principals stated they had sufficient support to prepare for the mid-year conference. Nearly 80% of respondents stated that their evaluator provided them with additional support when necessary during the mid-year conference process. Principals in low-poverty and high-performing districts were more likely to report having sufficient support available to them during this stage.

The case study interviews provided an opportunity to examine the mid-year conferences in greater depth from the perspective of both principal and evaluator. In these districts, the mid-year conferences were typically held between March and April. Preceding the conferences, the evaluators sent out guidance to the principals on what to bring and what questions would be asked. Principals brought their portfolios/evidence collections, which could include stakeholder survey results (when available), interim student achievement and growth data, and evidence of progress on their professional growth goals.

Conferences in the case study districts were guided by the evaluator's questions, which, in turn, were influenced by the questions the MDE suggested be used for the mid-year conferences. Principals responded with a presentation and conversation about their evidence. Evaluators typically asked clarifying questions and then provided feedback to principals on how they were doing in reaching their goals. At this stage in the evaluation process, results were not quantified. One interviewed evaluator said s/he focused

on strengths and weaknesses, first reinforcing accomplishments and then engaging in a dialogue with a principal on how to accelerate progress on their professional goals or student achievement outcomes when evidence didn't support enough growth.

Variations in How the Case Study Districts Conducted the Mid-Year Conferences

One of the districts held two “pre-meetings” with each principal before the mid-year conference. In the first meeting principals brought their student achievement data and in the second meeting they brought their evidence of progress toward their professional goals. The evaluator reviewed the evidence and offered feedback on its appropriateness.

We had one meeting where we talked specifically about the professional growth goals that the principals had and what evidence they've been collecting in that area. And at that point, I said, “Okay, well, that's the evidence you've collected so far between now and March, when we have our midyear conference. We agreed to any different type of evidence that they should collect between then and the midyear. And then in January, we had another similar meeting, except this time we were talking about their school improvement plan and the evidence they've been collecting in and around that.

In another variation, the evaluator shared promising leadership and instructional practices across the district. While conducting the mid-year conferences, the evaluator noted principals' promising practices.

They showed me evidence on each goal. I asked a lot of questions, kind of probing questions or have you thought of this, and one—one opportunity I always use in mid-year is I say I get the pleasure of meeting with every single principal and I get to see all these great things that they're doing or this particular tool they're using or how they analyze the data. So I ask, “Do you mind if I share some things I saw from other principals that maybe you could put in your tool box?” and they're always very, very open to that.

c) Preparation for the End-Of-Year Conference—Variations Between Principals and Evaluators

The top three activities of a large majority of principals during the months following the mid-year conference were

1. Communicating with the evaluator about the expectations for the end-of-year conference
2. Collecting and organizing evidence per discussion in the goal-setting conference and the mid-year conference
3. Collecting and organizing student assessment results.

According to principals' survey responses, after completing their mid-year conference, 25% of principals performed three to four actions from the list provided in Table 7,

while nearly half of principals engaged in seven to nine follow-up activities. Principals in districts with less ethnically diverse student bodies as well as those in low-performing and mid-size districts, on average, engaged in fewer activities in preparation for the end-of-year conference than other districts.

Evaluator survey responses confirm these findings. However, there again is a difference in reporting between principals and evaluators on the items of evaluator feedback, school visits, and observations. Although only about 7 in 10 principals reported that the evaluator offered feedback on their performance, professional growth, and action plan implementation, 13 out of 14 superintendents reported providing such feedback. An even greater disparity occurs in comparing principals' and superintendents' reporting about evaluator school visits and observations. Although 12 of the 14 evaluators said they conducted school visits and observations in preparation for the end-of-year conference, only 5 out of every 10 principals concur with this statement (see Table 7). Evaluators in non-rural districts were notably less likely to engage in this activity.

TABLE 7. Actions Principals Took Following the Mid-year Conference

Mid-Year Conference Follow Up Actions	Percent of Principals who Engaged in this Activity (N=48)	Percent of Evaluators who Report that Principals Engaged in this Activity (N=14)
The Evaluator Communicated the Expectations for the End-Of-Year Conference to the Principal	75.0%	100%
Principal Collected and Organized Evidence as Discussed During the Goal-Setting Conference and the Mid-Year Conference	72.9%	100%
Principal Collected and Organized Student Assessment Results as they Became Available for Demonstrating Progress Made on School Improvement Goals	68.8%	100%
Principal Implemented Action Plans for Professional Growth	66.7%	93%
The Evaluator Provided Feedback on Principal's Performance, Professional Growth, Action Plan Implementation, and Evidence Collection	66.7%	93%
Principal Implemented Action Plans for the School Improvement Goals	62.5%	86%
The Evaluator Conducted School Visits And Observations	52.1%	86%
Principal Completed Another Self-Assessment in Preparation for the End-of-Year Conference	47.9%	71%
The Evaluator Helped Principal Plan for the End-of-Year Conference	45.8%	86%

Variations in types of data discussed during the end-of-the-year summative conferences

There are substantial variations in the type of data that evaluators and principals discussed during the end-of-year summative conferences. Superintendent tenure in the district, urbanicity, student performance and socioeconomic status appear to be factors in these variations. Specifically:

Principals in districts with evaluators with longer tenures were more likely to discuss evidence of progress and outcomes of their professional growth goals as well as parent feedback; they were less likely to discuss student growth data and student feedback.

Principals in rural districts were less likely to discuss student feedback, while principals in low-performing districts were less likely to discuss student achievement data but more likely to discuss parent feedback.

Principals in low-poverty districts were less likely to discuss parent feedback and mid-year conference results and feedback. They were more likely, however, to discuss student feedback and documentation of professional development completed.

Although less than half (45.8%) of principals noted that their evaluators helped them plan for the end-of-year conference, more than three-fourths (77%) reported that their evaluators provided them with sufficient support to improve their professional practice prior to the end-of-year conference.

Supports for Principals

The case study interviews explored the kinds of supports provided by evaluators for principals to achieve both their professional and school improvement goals between the mid-year and end-of-year conferences.

District supports included providing *financial resources* to purchase instructional programs and their associated assessments, and approving the time principals needed to implement these programs, including workshops on specific instructional strategies (such as adopting READ 180, or developing education learning objectives [ELOs] and their associated formative assessments.)

They've gotten the financial support they need to carry out some of their ideas. We talked about the ELO's and attending those workshops. We talked about defining the idea that our elementary principal is going to change the Read Well by Third Grade plan. All those things cost money. They take time to review, they take time to change, they take time to develop. So the biggest thing that I feel like I've provided them between now and then is a clear vision that I'm going to support them in carrying out their plan if I agree with it. And to be honest with you, I think everything they presented to me is sound in terms of improving learning. It's research driven.

In terms of providing *learning opportunities* for principals, districts used monthly administrative meetings to further develop principals' skills in the common areas that principals wanted and evaluators thought they needed. Some of these included deepening

conversations on what the performance measures looked like in practice, revising rubrics to reflect the districts expectations for principals, training in cognitive coaching to help principals evaluate teachers, providing software to help principals undertake and give prompt feedback to teachers after their classroom observations, and supporting principals and their teachers to be culturally sensitive in terms of their instructional strategies. Quotes from three evaluators provide further context.

From mid-year until summative they [the principals]... they're still executing their school improvement plans, any professional growth plans that they have. So they're still working it right till the end of the school year and then they're turning their attention to analyzing the data. So did it make a difference? Did our action plans move the bar? The whole month of June we're focused on data analysis and then the other thing we were focused on, because we're trying to live our mission and achieve our vision, is creating a culture of innovation. So we had a two-day workshop of how to help principals create a culture of innovation that allows people to be creative and to do things differently, not always having to do things the way they've always done it.

Always tie everything back to your strategic goals, your mission, and your vision and then people will start seeing how it's all connected. That's really the key. We just had a brainstorming session on our continuous improvement journey. We meet for an hour and a half to two hours and that is just completely a learning environment where we're working on our continuous improvement efforts and we are collaborating and coming to consensus around—and using the PDA [Plan, Do, Act] process to really problem solve the issues in our district. So that's a road map for this year of where we're going.

There are collaborative networks among the principals for colleagues to sit down and talk about how they're doing on it. There's mentoring for new principals that come into the system to get up to speed on the system. Any time they have a question or they want some feedback or they want some group discussion around that, I've got that built into principal meetings.

Time spent on supporting principals also was an important resource. In practice, this means that most evaluators increased the time they spent in schools, including strategic observations of principal's leadership and classroom learning. The evaluator's commitment to visit every school and each and every classroom emphasized the importance of student learning in the eyes of the superintendent. The strategy always provided feedback for principals to improve their practice. Quotes from two evaluators provide further context.

So now I schedule a minimum of nine observations, and it's way more than that, but those are the scheduled ones that happen on an annual basis for each principal... I go and watch a PLC, a staff meeting, and a leadership team meeting, and I'm learning a ton from those, just how they interact with staff and what's their priorities, how they're pushing their school improvement plan forward. I mean, that's huge.

This year my goal is to visit every school and every classroom. So just being in the building so much more this year and seeing what's going on, because you know leaders impact the culture. I've attended some staff meetings, again, observational, been in the buildings on their PLC

¹⁹ Cognitive coaching is based on a set of principles that foster practitioners' cognitive development and autonomy that was developed by Arthur L. Costa and Robert J. Garmston. To learn more see *Cognitive Coaching: A Foundation for Renaissance Schools*, Norwood Massachusetts: Christopher-Gordon Publishers, 2002.

Wednesdays to see how they run—those kinds of things so mostly observational and, of course, what I learned through our meetings because our meetings are not really typically sit and get—they're problem solving, they're collaborating, so I'm observing their level of engagement and their level of involvement.

Principals primarily reported support for *improving instruction*. This could include additional financial resources to adopt or expand a curriculum program or providing time to teachers to align the reading curriculum with the Common Core State Standards. It might also include having an instructional coach or learning specialist in the school or having access to central office expertise on assessment, equity, or teaching and learning.

Other principals reported opportunities for *personal coaching*. For example, some evaluators dedicated a monthly administrators meeting to go over school issues or provide training in the cognitive coaching technique.¹⁹ Principals also benefitted from having individual support at the school to get guidance on making better use of time, cognitive coaching, and regular feedback.

The supports are almost built into my professional goals. For example, I have meeting time with our human resources director to do some cognitive coaching; not only is the notion of the meeting built into the goal, but now my schedule is set. So I think that those pieces that seemed to kind of need additional framing or additional resources have been put in place. I feel like I have everything I need to move in that direction.

This year for the first time, I have set up monthly meetings with my evaluator and that has been huge in terms of getting feedback. I ask him to tell me where you see my weaknesses and what do you think I'm missing?

d) End-of-Year Conferences—Variations Between Principals and Evaluators

According to the survey responses, the end-of-year summative conferences took place at different times over the course of the summer and early fall. Although some districts completed their end-of-year conferences before or immediately after the school year ended, other districts continued holding summative conferences with principals in September and October. Despite the variation in end-of-year conference times, 81% of the principals agreed that end-of-year conferences occurred at optimal times for them.

Evaluators' survey responses on the timing of end-of-year conferences, however, differed from those of principals. Although 69% of evaluators thought that the end-of-year conferences occurred at an optimal time for them, only half of evaluators thought that the timing of end-of-year conferences was optimal for principals.

The case study interviews provided additional insights about the timing of the end-of-year conferences. Evaluators saw advantages in either completing principals' professional growth goals evaluations and stakeholder survey results evaluations by June (and finishing the process with a final summative rating in the fall when the MCA data became available) or completing the entire summative review in June using interim student achievement data.

I met with each principal one on one. That's when they brought all their evidence to show where they believed they were in the ratings and to help me create that summative rating for each of those five performance measures, because like I said to them, "I can't possibly know all the great things you do. I can know the things you tell me or the things that I observe. So don't be shy about it. If you feel like you deserve a Distinguished in a certain category you need to tell me why.

Evaluators noted that either completing the full summative review in May or June or partially completing it gave principals time in the summer to work on their professional growth goals. As one evaluator explained:

But to do it in June, we need look at a much broader variety of achievement data. In my mind the optimal time to get these evaluations, summative evaluations, completed would've been in June. So they would've had the summer to reflect on those, to think about how they're going to incorporate improvement strategies into their next evaluation cycle and their self-evaluation. So I feel we have to have that [summative review] in May because then it gives us two months as a team to decide what type of activities we need to implement or we need to attend to make our leadership skills a little sharper. And I would try to utilize as much professional development dollars as I could through the Title II program.

Evaluators also talked about the timing of the stakeholder surveys for use in the summative review. If done too early in the year, parents and teachers may not be familiar enough with the principals' leadership skills. If done too late in the year, there isn't enough time to reflect on the findings and take actions to improve stakeholders' perceptions. One evaluator talked about giving the survey early in the academic year and then again late in the academic year to identify a principal's professional growth. In another district, evaluators thought administering the stakeholder surveys in February or March would allow enough time for principals to reflect on the results and still take actions during that school year.

Principals tended to agree that an optimal time of year for the summative review would be in June. Having the summative review in advance of the next year's goal setting made a lot of sense to principals because it meant having the summer to engage in related professional learning opportunities. Secondary school principals, unfortunately, depend on the Multiple Measurement Ratings (MMR) data to know if they met their goal on increasing graduation rates. But proficiency, student growth, and achievement gap reduction can be estimated by some interim measures that districts use that sometimes also predict student scores on the MCA.

It would have been better if we'd been able to meet in June, because then you don't have as many... interruptions isn't the word...maybe distractions would be a better word. Trying to get things off the ground in the fall is difficult at best, plus we've got so many other projects that need to get done on top of that. It would have been much better off if we had met in June.

It [summative review] should take place before June 30. The whole process should take place between July 1 and June 30 of the next year...I'm already on the first lap of '13-'14, developing our stakeholder goals with our staff and doing our academic goals. And I'm still having to pay attention to last year, to provide the data and to meet to get the evaluation. When it comes to June 30, I want to cut last year off and focus everything on this year.

Sources of data used for the end-of-year conference

The top three data points that principals discussed with their evaluators during the end-of-year summative conferences were (1) teacher feedback (77%), (2) student achievement data (77%), and (3) principal self-assessment results (72.9%). The least frequently discussed items were student and parent feedback as well as evaluators' notes from interactions with teachers and other members of the school community (see Table 8).

TABLE 8. Data Principals Discussed with their Evaluators During the End-of-year Conference

Data Points	Number of Respondents	Percent
Teacher Feedback	37	77.1
Student Achievement Data	37	77.1
My Principal Self-Assessment Results	35	72.9
Student Growth Data	33	68.8
Documentation of the Professional Development Completed During the Year Aligned to my Professional Growth and School Improvement Goals	28	58.3
Indirect Observations of my Practice (Such as Classroom Instruction, Grade Level or Department Planning Meetings, Peer Coaching Sessions, Etc.)	23	47.9
Mid-Year Review Results and Feedback	21	43.8
Feedback from Non-Teaching School Staff	19	39.6
Direct Observations of my Practice	19	39.6
Student Feedback	17	35.4
Parent Feedback	17	35.4
Evaluators' Notes from Interactions with Teachers	17	35.4
Evaluators' Notes from Interactions with District Staff	10	20.8
Evaluators' Notes from Interactions with Other School Community Members	7	14.6

According to their survey responses, during the end-of-year conference principals discussed an average of seven points of evidence of their progress on professional growth and school improvement goals. Specifically, 54.2% of respondents reported discussing between 6 and 10 points of evidence with their evaluators. About one-fifth of respondents (18.8%), however, discussed three or fewer sources of evidence. Principals in low-poverty and non-rural districts, on average, discussed fewer data points with their evaluators during end-of-year conferences.

In the case study interviews, both evaluators and principals reported using student achievement results as evidence of progress on school improvement goals, principals' portfolios for their evidence of progress on professional goals, evaluators' interactions with principals, and stakeholder survey data. The example evaluation model also requires using measures for the high school graduation rate, student proficiency, student growth data, and achievement gap reduction.

Case Study Evaluators' Perspectives on Evidence and Data Sources

The interviewed evaluators reported that they had used interim measures for many years to place students into instructional groupings, revise curricula, and monitor student progress. In this study, some of these measures were associated with and aligned to a particular curriculum that the district has adopted. Districts now also use these interim student assessment results to gauge principals' progress on their school improvement goals, supplemented later in the year by the MCA results. Some of these assessments, when administered at least two times a year, provide student achievement growth data as well as progress monitoring.

All four case study districts used different stakeholder surveys and administered them at different times of the year. One district had developed teacher, parent, and student surveys the year before so it had baseline data for year-to-year comparisons. Two districts used stakeholder surveys recommended by the MDE, and another used a commercially available survey aligned to the Minnesota Performance Measures and Indicators. Evaluators reported that stakeholder survey results held significant sway in evaluating principals on the five performance measures.

These evaluators also reported that principals' questions, participation, and contributions in cabinet, school board, or administrative meetings also factored into their evaluations of principals.

I think between the reflective piece with their self-assessment and the rich dialogue that brings you [together] in the fall, tying the school improvement plan in, tying in academic data and stakeholder perception data, it's really a well-rounded package. I feel like I have a lot of data to support my ratings and to support my opportunities for [suggested principal] improvement. They're not just getting plucked out of the air.

Finally, the evaluators talked about other observations of principal practices that contributed to the ratings. These included how well they communicated with their staff and parents, whether their budgets were underspent or overspent, how data were being used at staff meetings, and whether principals were doing effective classroom observations on standards-based instruction.

Case Study Principals' Perspectives on Evidence and Data Sources

In their responses during the case study interviews, principals were much more focused on how these data sources could be used to improve student achievement in their schools. Both proficiency and growth data provided by interim assessments were crucial to identifying appropriate student interventions, developing school improvement plans and then setting the following year's school improvement goals. These, in turn, informed the principal's school improvement goals.

Table 9 shows the sources of data that principals and evaluators cited during the interviews and how these sources varied by the different school levels.

Until the MCA and MMR data became available, these interim measures were used to monitor student progress toward principals' school improvement goals. These measures also provided valuable feedback during the mid-year conferences.

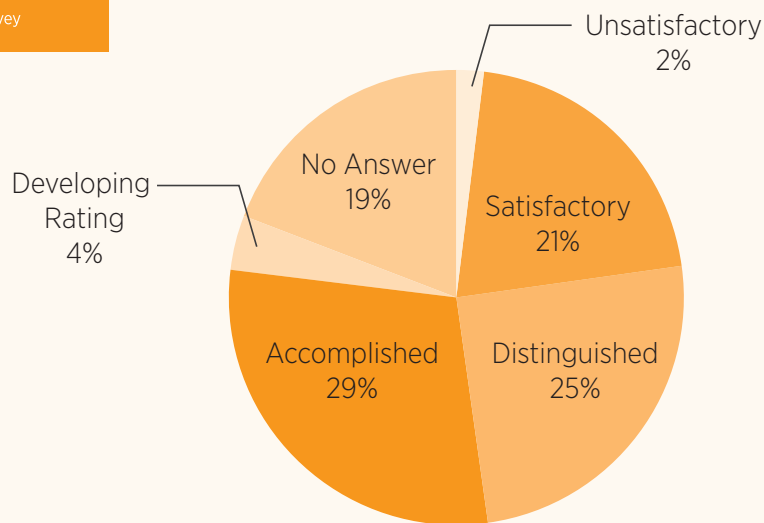
TABLE 9. Data Sources Used in Districts as Reported by Evaluators and Principals in the Interviews

Type of Data	Source of Data
Stakeholder Surveys	<ul style="list-style-type: none"> • Five Essentials (University of Chicago) (survey of teachers) • Minnesota Principals Assessment Survey (survey of teachers, although a parent survey exists) • Delaware Performance Appraisal System (DPAS II) (survey of teachers) • District developed teacher, parent, and student surveys • District developed parent survey
Elementary Schools	<ul style="list-style-type: none"> • DIBELS as an early reading inventory (Dynamic Indicators of Basic Early Literacy Skills) • Northwest Educational Associates Measures of Academic Progress (NWEA MAP) (predicts proficiency and measures growth in academic subjects) • Accelerated Reader mathematics and reading assessments • AIMSWeb in reading and mathematics (available for grades K-12) • Optional Local Purpose Assessment (OLPA) (Available for grades 3-8) • MCA reading proficiency in grades 3-8 • MCA mathematics proficiency in grades 3-8 • Multiple Measures Rating (MMR) (proficiency, student growth, achievement gap reduction, graduation rate)
Middle Schools	<ul style="list-style-type: none"> • READ 180 (a curriculum with associated student progress assessments) • Northwest Educational Associates Measures of Academic Progress (NWEA MAP) • Number of discipline referrals by ethnicity MCA (grades 3-8) • Multiple Measures Rating (MMR) (proficiency, student growth, achievement gap reduction, graduation rate)
High Schools	<ul style="list-style-type: none"> • ACT Plan and Explore data • Advanced Placement (AP) test scores results • District developed common assessments • Passing rates (credit acquisition between 9th and 12th grade) • MCA grade 10 reading, grade 11 mathematics • Multiple Measures Rating (MMR) (proficiency, student growth, achievement gap reduction, graduation rate)

Summative ratings could be completed once the MCA and MMR data were received in the fall. Principals reported variation in the summative ratings they received (see Figure 9, below). The majority (54.2%) received Distinguished or Accomplished summative ratings. About a fifth of principals (20.8%) were rated Satisfactory. Only four percent of principals reported having Developing added to their rating. Almost a fifth of the principals (19%) did not answer the question.

FIGURE 9.
Distribution of Summative
Ratings among Principals

SOURCE: Principals' Survey



e) Evidence Used for the Quantitative Rating by the Case Study Evaluators

During the case study interviews, all of the evaluators reported they had developed a quantitative rating process to calculate and assign ratings to principals. All four case study districts used the recommended weights of 15% for stakeholder surveys, 35% for student achievement and growth, and 50% for the supervisory rating.

Stakeholder Survey

Because the four districts used four different stakeholder surveys, each had developed a different method of creating ratings to quantify the results. Some districts disaggregated survey data into the five performance areas and came up with an average rating that was compared to a target (that is, 90% of parents will be satisfied with the school). If a principal missed a target but was close to it, then partial credit could be given. For example, a district that gave three surveys (teachers, parents, and students) could assign a weight of 5% to each survey. If the principal met the 90% satisfaction target on two of them, then he or

she would get 10%. If the remaining survey had an 87% satisfaction rate, the principal might be given 4% partial credit for a total of 14% of the 15%. For surveys that were normed, percentile rankings using quartiles were used to assign one of four performance levels to principals for the survey component of the rating.

In another case study district, points on the stakeholder survey were based on a principal's reflection on the results of the stakeholder survey. Principals who analyzed their survey results and made planned or real changes in their practice, the evaluator awarded points for the stakeholder survey. The district that scored the principals' reflection on the survey results looked for deep insight from the principals.

Student Achievement

Evaluators also reported how they rated the school outcomes. Because of the diversity of goals and how targets were set, there was variability in how districts calculated ratings for these outcomes. This third component of the principal evaluation included student achievement and growth scores as well as principals' progress on their professional and personal growth goals, which across the four case study districts could vary from as few as three goals to as many as nine for individual principals.

Then we went to the school outcomes portion, and for that, principals basically presented us with documentation, graphs, whatever they had. We had one individual that brought us a PowerPoint presentation and basically walked us through it. "Here are each of my goals. Here's how well I did in terms of reaching those goals. If there were reasons why I didn't reach them, I want to reflect for you why I didn't." And then they did a similar situation with their individual professional improvement goals and we also looked at those. And again, they gave us the evidence that would show to what extent they were able to achieve those. And then we reviewed all the components that they identified with all the data they provided, and again, ranked that section. And then in the end, then we went back to [the MDE] template...and inserted their proficiency rating.

So here's an example of the gap goal. The gap goals—we have two of them—were worth 15 points. So one was worth seven and a half and the other seven and a half, but together basically we said at 0 would be the gap widened, 5 the gap was maintained, 10 there was some progress in closing the gap or 15 you attained your goal. So that's how we did it in between there and then we were able to mathematically come up with a score.

The MCA results were the primary source of data for measuring student achievement (a minimum of 35% of the weighting). Often these results included student proficiency and growth scores in reading and mathematics, closing the achievement gap, graduation rate (for high schools), and sometimes a district-specific goal. Typically, different schools set different targets, but targets were designed to assure all schools met the NCLB Adequate Yearly Progress goals by 2020. The one district that didn't wait for the MCA data to arrive used interim achievement and growth measures to evaluate principals' on their schools' student achievement.

²⁰ In holistic scoring a score is determined by an overall impression of the evidence. Because this district used a narrative by the principal for the self-assessment, the evaluator assigned a holistic score for the appropriate performance level and then used it in the calculation of the final supervisors rating.

Evaluator's Ratings

The case study evaluators used multiple sources to rate the principals' progress on professional growth goals and school improvement goals, as well as the evaluators' own ratings on the five performance measures (worth 50% of the total score). These sources included the principals' self-ratings, evidence principals presented, and evaluators' observations and interactions with principals during the year. Stakeholder survey results that had a leadership component aligned to the five performance standards might also be used in calculating these ratings. In one district, principals could select one item off a stakeholder survey, set a target for it and use that as another professional growth goal. The principal's school improvement goals also were in this category. Points were awarded, to the extent the goals were measurable and progress made on them.

f) How Case Study Evaluators Calculated the Final Rating

The case study districts typically developed a spreadsheet (or purchased a vendor's spreadsheet as part of a package) to house individual ratings and create a final summative rating. Evaluators reported that they asked each principal to come to the end-of-year conference with their own ratings and evidence on the five performance measures (for the evaluator's rating). Before the conference, evaluators also had independently assigned principals' ratings on the five performance measures. During the conference the principal and evaluator discussed the evidence marshalled to support the ratings. In most cases the principal and evaluator had come up with the same rating.

In one case study district, the principals created narratives, and the final rating was done holistically.²⁰ Another district scored the performance measures holistically without the narrative. In the other two case study districts, quantitative ratings were assigned based on the descriptors in the performance levels.

I'd look at the rubric and I'd say, "Holistically, what does it look like you're doing—how you're doing on this rubric? Okay, then holistically, it looks like you're Proficient. Then once I put all those Proficients in—I didn't average them, per se, but once I put all those Proficients and Satisfactorys and Distinguisheds together, I'd look holistically at that and say, Okay, this is how you're doing for the overall supervisor rating."

g) Perceptions of the Final Ratings

According to Table 10, the large majority (more than 78%) of principals agreed with their performance measure ratings. However, only 64.3% of principals agreed with their rating for stakeholder feedback. Although most evaluators thought that their ratings for performance measures 1 through 5 accurately captured principals' performance, fewer of them—5 and 9 out of 14 principal and evaluator respondents respectively—agreed that school performance data and stakeholder feedback were accurate measures of principal performance.

In Figures 10a-c, we examine variations in perceptions across different district contexts. The ratings for stakeholder feedback and school outcomes showed the greatest variation in the level of principals' agreement with the final summative rating across all the different contexts.

TABLE 10. Principals' Perception of the Summative Ratings

Performance Measures	Disagree	Neither Agree nor disagree	Agree	Number Reporting they did not Receive a Rating
Rating for Performance Measure #1: Establish a vision and mission focused on shared goals and high expectations	7.10%	9.50%	83.40%	1
Rating for Performance Measure #2: Provides instructional leadership for high student academic performance	7%	11.60%	81.40%	
Rating for Performance Measure #3: Manages human resources for quality instruction and professional growth	7.30%	14.60%	78.10%	1
Rating for Performance Measure #4: Builds professional and ethical relationships through collaboration and effective communication	7%	9.30%	83.70%	
Rating for Performance Measure #5: Strategically manages resources for systemic performance accountability	4.90%	17.10%	78.00%	2
Rating for Stakeholder Feedback	9.50%	26.20%	64.30%	1
Rating for School Outcomes	9.30%	9.30%	81.40%	
Final Summative Rating	7.50%	12.50%	80%	

Figure 10a shows that principals from larger districts, on average, were more likely than principals from smaller districts to agree with their summative ratings, with the single exception of the rating for stakeholder feedback.

Although principals in low-performing districts, on average, are less likely to agree with all the ratings, they are least likely to agree with their ratings for stakeholder feedback, school outcomes, and the final summative rating, as shown in Figure 10b.

Principals in districts with superintendents with shorter tenure, on average, are less likely to agree with the stakeholder feedback rating, and principals in larger and high- or some-poverty districts are less likely to agree with their rating, as shown in Figure 10c.

In case study interviews, the principals indicated that they believed the evaluation process was fair and that they agreed with the ratings and how the ratings were determined. But a few principals in the case study districts reported a bit more skepticism than did their evaluators in how the ratings were determined. In a few instances, principals didn't agree with their ratings—although most went along with them because of the low stakes nature of the pilot and a commitment from the evaluator that changes in the process would be made for the following year. One principal commented that a quantitative number didn't tell him how to improve. Quotes from three principals provide further context.

I think both evaluators looked at the information that I had provided for them, and then they also looked at the big conversation that we had together on the reflections. Then they made the final determination on the rating at that time. Did I think it was fair? Absolutely!

This rubric was developed with the help of some building principals and district administration... And then the superintendent put numbers or ratings based on that. So that's how it went for each of the [performance measure] areas...Oh, absolutely these are good descriptors. I think it's if we're self rating ourselves, we need to know what to base that on so we can read each of these and reflect on where we see ourselves. So I think that this was helpful.

For me, the final evaluator's rating number, we didn't agree, but then I told you earlier that there were some things that went on in my school. When I think about the stakeholder feedback from teachers—we had some things that happened last year, and we had to create a new scenario for us this year. I'd like to think that if we had done that survey later, it still might not have been what I wanted, but it would have been better, because we had worked through some things here at our school...The rating was fair based on the data we had in front of us. Unfair because a lot of things were out of my control.

Evaluators reported experiencing challenges in assigning ratings for stakeholder feedback as well as school outcomes (see Figure 11). Specifically, more than two thirds and over a half of respondents experienced challenges assigning school outcome and stakeholder feedback ratings, respectively. Nonetheless, 100 percent of evaluators agreed that the summative ratings they assigned to principals they evaluated accurately captured their performance.

FIGURE 10, A-C Level of Principals' Agreement with their Summative etc.
 (Note: The higher score denotes a higher level of agreement)

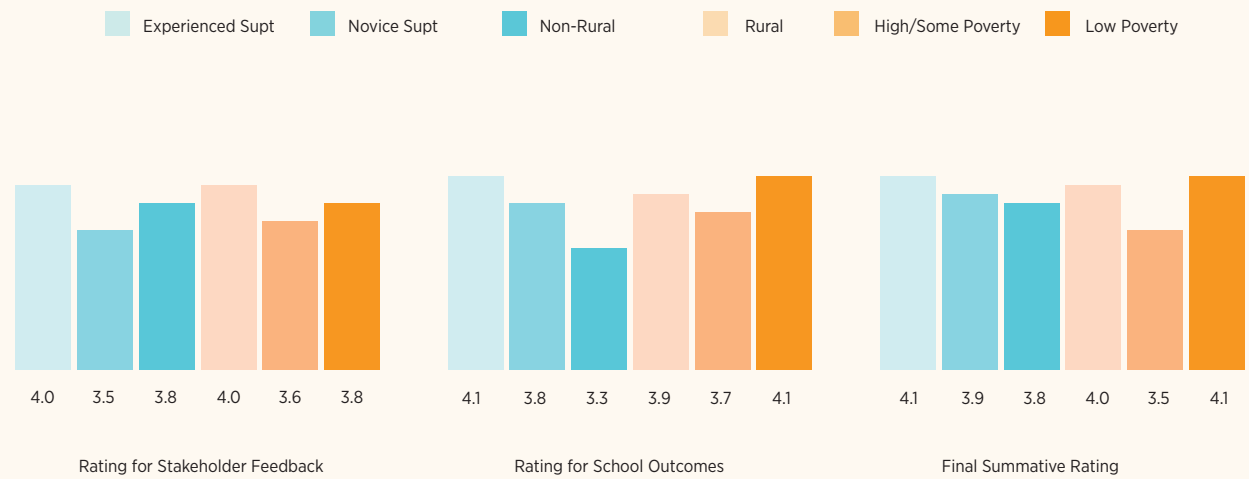
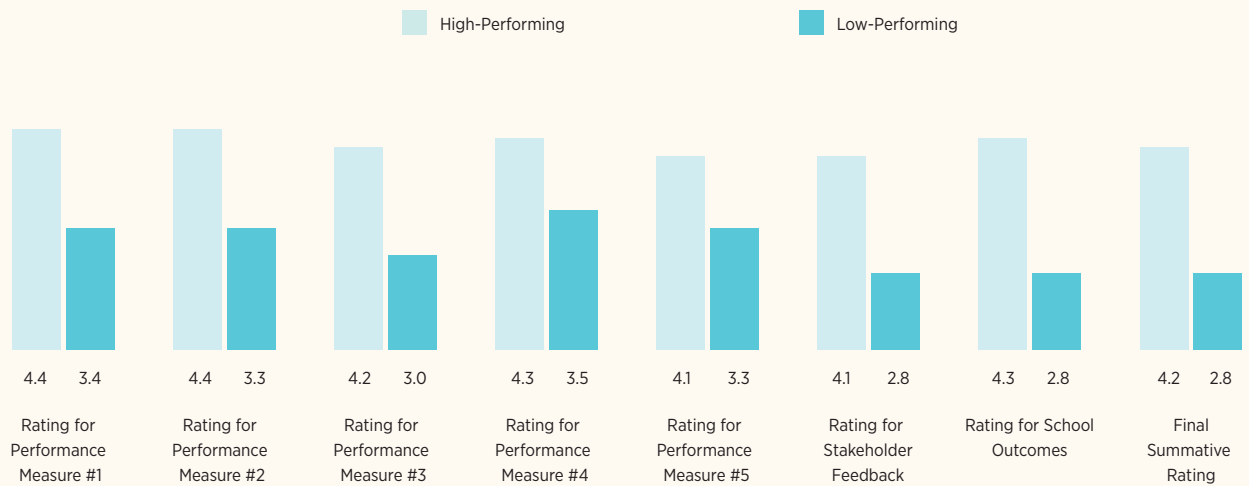
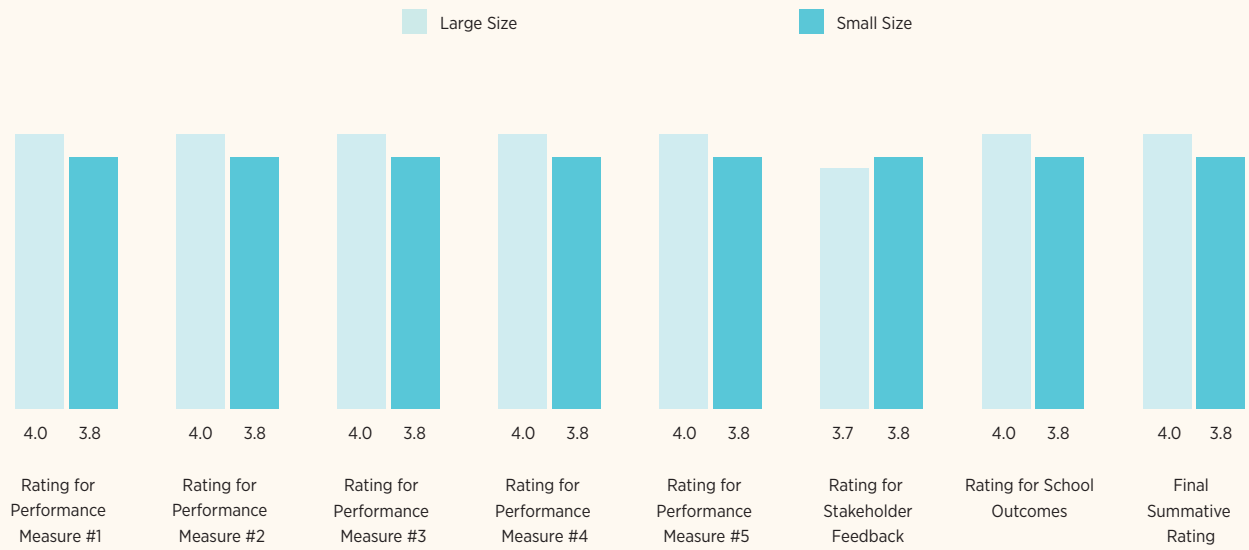


FIGURE 11.

Number of Evaluators who Experienced Some Degree of Challenge Assigning each Summative Rating to Principals (N=13)



38.5% Rating for performance measure #1: Establishes a vision and mission focused on shared goals and high expectations



30.8% Rating for performance measure #2: Provides instructional leadership for high student academic performance



30.8% Rating for performance measure #3: Manages human resources for quality instruction and professional growth



30.8% Rating for performance measure #4: Builds professional and ethical relationships through collaboration and effective communication



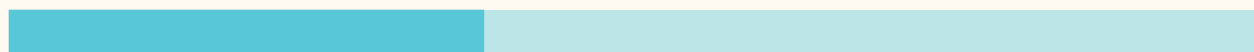
30.8% Rating for performance measure #5: Strategically manages resources for systematic performance accountability



53.8% Rating for stakeholder feedback



69.2% Rating for school outcomes



38.5% Summative rating

**h) Case Study Evaluators/
Reflections on
the End-of-Year
Conference Process**

In the case study interviews, evaluators were asked about any difficulties they faced in either preparing for or taking part in the end-of-year conference. One evaluator noted a need for more guidance in constructing questions to pose to the principals. Another evaluator identified the delay in getting the final MCA results to finish the end-of-year conference, which was to be a principal's annual employment evaluation and contract renewal. This same evaluator talked about the change in the MCA reading assessment and the inability to calculate comparative student proficiency and growth ratings from the previous year. A third evaluator felt evaluators had the "easier end" of the evaluation process because the principals were responsible for collecting their evidence and evaluators had to merely "listen and ask really great questions." Yet, by asking these questions principals began to see the similarity between their evaluation and the evaluations they were doing with their teachers. "It was an 'ah ha' moment." Finally, the fourth case study evaluator noted that the principalship is a "nuanced profession" and worried about getting a principal's performance down to a single number. This evaluator counseled principals that it wasn't about their final number but identifying "some things in their practice that they could work on to make them a more effective and efficient principal."

3. What challenges were faced in implementing principal evaluation?

Challenges were identified throughout the study—both through the two surveys and the two rounds of case study interviews. Although some of these challenges were identified previously in this report, we summarize them here.

The primary challenge principals faced prior to the mid-year conference was collecting stakeholder feedback from parents, students, and teachers. Other challenges principals encountered during the mid-year conference preparatory period included finding time to organize evidence, make adequate progress on action plan(s), and reflect on individual professional growth goals. Interestingly, finding time to reflect on school improvement goals was less of a challenge for principal respondents than finding time to reflect on personal professional growth.

As shown in Table 11, when reflecting on the entire process—not just the mid-year conference—principals cited the collection of feedback from parents and non-teaching members of the school community as the most difficult element of the model to implement. Indeed, about a fourth of principals said that they never collected feedback from parents, students, and non-faculty members of the school community. In addition, 14 principals said they did not receive feedback from their evaluator that was specifically based on observations.

TABLE 11. Elements of Principal Evaluation that Principals Found Difficult to Implement

Elements of Principal Evaluation	Not Difficult	Somewhat Difficult	Difficult	Not Applicable (Selected if this element was not implemented in the district)
Feedback from Other School Community Members	28.0%	28.0%	44.0%	16
Parent Feedback	26.9%	38.5%	34.6%	15
Student Feedback	37.9%	34.5%	27.5%	13
Evidence Collection	43.9%	31.7%	24.4%	1
Review of Student Achievement Data	46.3%	39.0%	14.7%	1
Mid-Year Conference	53.7%	31.7%	14.6%	1
Evaluator's Feedback Based on Observations	57.1%	28.6%	14.3%	14
End-of-Year Conference	58.5%	29.3%	12.2%	
Teacher Feedback	55.0%	37.5%	7.5%	2
Self-Assessment	73.8%	19.0%	7.2%	
Goal Setting	71.4%	23.8%	4.8%	

**a) Challenges Reported
by the Case Study
Evaluators**

The four case study evaluators noted a number of challenges to implementing principal evaluation. These ranged from knowing whether they were implementing the model as intended, scheduling the conferences with principals, dealing with ineffective principals, the timing of receiving the student achievement and growth data, dealing with the “unwieldiness” of the written model materials, the need to be organized around aggregating the principal’s data, and finding the time to observe principals.

One evaluator wondered about doing the principal evaluation process “right.”

There’s so many ways that you can take this model, and how do you know you’re right? That was the thing because there is no real manual that you can go to and say this is the right way to do it because there really isn’t a wrong way to do this, either. So, I guess, all in all, no matter how you phrase that question, whether it’s mid-year, whether it’s end of the year, whether it’s the beginning of the year, is just a reaffirming that we’re doing this the right way, that we’re doing what we’re supposed to do.

Another evaluator was concerned about finding a good time to schedule the evaluation conferences so that it didn’t interfere with principals’ primary leadership responsibilities.

I think coordinating schedules, while it wasn’t insurmountable, that is a vexing piece in terms of just the number and the time, because we wanted to have deep conversations.

Dealing with an ineffective principal was a challenge for another evaluator, although the evaluation process made it easier to deal with principals who were not meeting expectations.

It’s always challenging when you’re dealing with an ineffective principal. Those are hard conversations. Those are courageous conversations. But no, as far as a process standpoint, it [the evaluation process] made it easier to deal with an ineffective principal. It, as I said earlier, it brought it to light. I mean I could tell she knew I knew. You know what I mean, because I was – oh, you can’t, you know, and then finally she just admitted [that someone else] had pulled all of this data for her. So it made it much easier to not hide behind jargon and stories about staff or students or parents so we really could keep focused on what matters.

Several of the evaluators noted that some of the model’s tools were “unwieldy.” This was especially true of the descriptors in the rubrics to identify the four levels of performance for the performance measures and their indicators.

That I think it was unwieldy. The model as written was unwieldy, and so we had to trim it down, make it more meaningful; I think that’s the biggest challenge that I faced.

The final rating system required evaluators to use a computer-based system for inputting the data into a spreadsheet, weighting the ratings, and producing a final rating for each principal. One very organized evaluator noted that initially it took looking across a lot of different district data sets to create an integrated spreadsheet to come up with a final rating for a principal. It also meant making time to undertake the evaluation system.

Just being organized. When you're doing something new and there's a lot of moving parts just being organized to do it and then devoting the time, reprioritizing something so that I truly was making this a really high priority part of my job.

The case study evaluators reported somewhat different perspectives on the time spent on observing principals. Some had always had frequent interactions with principals, visiting their schools, helping them problem-solve certain situations, observing their interactions and their contributions in administrative staff meetings (previously described). They were going to continue to do this. But several evaluators "upped the ante" and systematically scheduled visits to principals in their schools (observing particular events and situations) on an ongoing basis.

To do the level of observation that I think you need to do to make this a meaningful process, you need to set aside a significant chunk of time to be in the sites and doing observing and finding the right types of things to observe....I go and watch a PLC, a staff meeting, and a leadership team meeting, and I'm learning a ton from those, just how they interact with staff and how they set their priorities, how they're getting their mission out there, how they're pushing their school improvement plan forward. I mean, that's huge.

**b) Challenges Reported
by the Case Study
Principals**

In the case study interviews, the principals echoed many of the concerns of the evaluators. Principals also were unclear about the details of implementing the model, concerned about the availability of student achievement data, pressed for gathering and organizing the evidence on progress on their goals and performance levels, and worried about the time to get all of it done.

One principal wanted a more structured and well-defined evaluation process.

I think when it was first introduced there was a lot of gray area. In order to implement something like this and do it well you've got to eliminate all the gray. It's got to become very black and white, very structured. I know the complexity of doing that because we have so many different districts with so many different needs. But the structure itself should be more solid. I mean we had to basically hire somebody to come in and tell us this is how to do this, this is how to do that, and this is how to do the other when if things were just laid out a little differently I think it would have worked out a lot better. That was one frustration I had.

The delay and changes in testing data were universally criticized, but another challenge for principals was ensuring that the PLCs in their schools were looking at student data and making instructional adjustments to meet the school goals.

You know one of things that has come out of the whole end process is looking at the work with the PLCs, and having the dialogue with the groups, and taking a look at their data, and then coming back to them and asking those tough questions. Based on last year's reading data, What are you doing to address that? It's not about blaming, but, What are you going to do to address that?

Principals were challenged by collecting the evidence of their progress and skills and putting it all together in a timely format that could be shared with their evaluators.

Probably putting the pieces together. Obviously, the self-assessment is not hard to do, but gathering the information that's needed to set your goals, measure your goals throughout the process, and then report on that. And that, in some cases, I'm not the creator of the data, so I'm reliant on the person at the district level that is pulling the data together.

One principal in a non-Title I school felt that the accountability of student scores for the MCA was on the principal, who wasn't always supported by the teachers.

The math and reading tests for AYP purposes is only in Title I schools. We're not a Title I school. Why will the kids care? Why will my teachers, care? Some of them tell the kids, Don't worry about the test; it doesn't matter anyway. And now I am accountable but no one else is.

Like evaluators, principals were concerned about the amount of time they spent on the principal evaluation requirements—such as the self-evaluation, collecting evidence, and conversing and conferencing with the evaluator.

Time spent working on the self-evaluation, time spent gathering data, time spent meeting, time that you didn't—I mean going through it the first year you have to make room for it. While you did it, it took away from other things. When I did my self-evaluation and set myself up for my goals for this year, it didn't take near as much time because I knew what I was kind of doing where before it was reading through all the rubrics, trying to figure out how does that apply to me and how does that apply to my school, and how do I answer this.

As noted previously, principals reported challenges regarding the stakeholder survey data for several reasons. For one principal, the challenge involved what was happening in the school just prior to survey administration. Another principal worried about participation rates—if they were too low would the survey results be representative, thus credible and worth acting upon? Were there norms or comparable administrations of the survey that would provide benchmarks for target setting or year-to-year comparisons?

Finally, the burden of producing evidence was on the principal.

And the part that was challenging for me was not knowing what evidence the evaluator wanted presented and in what kind of depth. So that made it difficult to prepare. And when I visited with [the evaluator] about that last week at that meeting, she said her feeling was that principals were all along that continuum, and those that were involved in helping with the rubric had a distinct advantage because they knew exactly because they worked on it. So that was a first-time-through challenge. I feel now like I have the understanding, but I didn't. I kept re-reading that e-mail and principals were talking to each other. "What do you think this means?" So I know that the 10 of us that were not in that process of developing that, I wasn't the only one confused on that. So that was just part of the roll out. Now we know.

4. What Conditions Support Successful Implementation of Principal Evaluation?

Principals contended that the implementation of principal evaluation is most successful when it is primarily used as a professional growth opportunity, secondarily as an accountability system, and when it is implemented as a team effort between principals and evaluators.

Principals also said that the new evaluation process is unlikely to yield substantial value for principals' professional growth and school improvement if it is not a priority for district leadership or if evaluators are not trained to effectively facilitate the process.

Some principals also expressed concerns regarding inadequate or missed opportunities to reflect on data due to limited access to the data. They also emphasized variations in context from district to district (identified in their orientation trainings) or from school to school that need to be taken into account and planned for to ensure effective implementation of principal evaluation.

Beyond the principals' recommendations, a number of supportive conditions emerged from the case study interviews, some created by the state, others created by the district.

a) State Conditions

State conditions that supported implementation of the model included:

The state- and local-level education stakeholders' support for the example model.

Because a diverse working group developed the model, the state was able to create an example model that would work throughout Minnesota and have support of key stakeholders.

The local flexibility built into principal evaluation.

Evaluators appreciated the opportunity to make adaptations to the example model to fit their particular district contexts, priorities, and strategies. These included using a narrative for the self-assessment, trimming and tightening the performance measures rubric, developing their own surveys, and using interim measures to assess progress on the goals.

The effectiveness of the state orientation trainings for the example model.

Both evaluators and principals felt they had sufficient orientation to the example model to get started. Although districts still had many decisions to make, a model framework and a process with supporting materials had been put into place and shared with districts.

The example model's heavier emphasis on principals' professional growth versus a simple accountability model.

Throughout the case studies, principals indicated they valued the dialogue with their evaluators and the purposeful feedback they were getting on their goals. In addition, the performance measures and supporting rubrics gave principals more clarity on their role as principals and helped structure and support the work of their evaluators.

MDE's quick and effective response to pilot districts' questions and issues.

In addition to the orientation training, the MDE staff met with evaluators individually to answer any questions they might have. They also responded quickly to phone calls or e-mails from the pilot districts as issues emerged. One evaluator was grateful for the "just-in-time" e-mail reminders (with pertinent materials attached) to undertake the model components.

b) District Conditions

District conditions that supported implementation of principal evaluation included a commitment to continuous improvement, capacity to support professional growth, and the use of proven strategies/support. These are briefly described below.

Districts had a school board that supported the piloting of the new principal evaluation process in their district.

In the case study districts, the school boards approved their districts participation in the pilot and were kept up-to-date on the district's evaluation activities by the superintendent.

District leaders believed in the potential efficacy of the principal evaluation process in improving school leadership.

For whatever reasons the districts volunteered to join the pilot, they all saw it as an opportunity to strengthen instructional leadership in their district.

District leaders fostered an articulated alignment between the district mission, vision, district goals, strategic plan, school improvement plans and principals' goals.

Evaluators often had a focus on streamlining the principal evaluation process by aligning it with the district's goals and strategic plan. This alignment allowed principals to piggy-back their school improvement goals on their schools goals and set their professional growth goals to support their school improvement goals.

Districts had a history and culture of school improvement and accountability.

For many years schools in Minnesota had been making annual school improvement plans and had district data systems to support them. Principals had prior experience with progress monitoring and meeting school improvement goals. In some districts the school improvement plans were presented to the school board, fostering accountability for reaching their goals.

Districts were committed to using stakeholder surveys as a key source of "customer satisfaction" evidence.

Most of the case study districts used the results of the stakeholder surveys as important sources of evidence of how the principals were being viewed in their school and community. One of the case study districts had administered the surveys the year before and could use the surveys as a key resource in principals' goal setting. Other districts used the results of evidence of principals' success on the five performance measures for feedback and ratings.

Districts had evaluators who were skilled at developing collegial and trusting relationships with principals.

Evaluators were optimistic about results when they and principals worked as a team, shared honest feedback (even when the news wasn't good), and brainstormed together to solve schooling issues. Several of the case study districts had a culture of continuous improvement where they "fixed the problem and not the blame."

Districts had evaluators who helped streamline the model and develop integrated technology to track principals' progress and calculate a final rating.

Two of the case study districts revised the rubrics to align with their school improvement strategies. This made it easier and more accurate to assign a performance level to a principal. Evaluators also developed spreadsheets to house all of the principal's data in one place—making final rating calculations possible.

Districts had highly qualified leaders in key roles to support principals.

These individuals were often experts on district data and assessment, equity, teaching and learning, and human resources. These leaders provided supports to principals, linked them to research-based practices, coached them, and provided feedback when requested. In one district, principals presented their goals to a panel of these leaders, so the district had a heads-up on what these principals would need in terms of district supports to meet their goals.

Districts where both evaluators and principals re-prioritized their time to invest in the principal evaluation process.

Principals talked about the time they spent in implementing their plans to achieve their goals and collecting supporting evidence. Evaluators talked about having many more interactions with principals and using administrative meetings to help principals deepen their understanding of the performance measures. Both evaluators and principals agreed this was worth the time they were putting into it.

Districts realigned the agendas of their monthly administrative meetings with principals to support the evaluation process.

Case study districts used their biweekly or monthly principals' administrative meetings to help principals deepen their understanding of the example model and the performance measures. This included rewriting the rubric descriptors using language that resonated with the principals; identifying the most important performance measures to assess; training principals in SMART goals; refreshing principals' skills in creating a school's vision, mission and goals; and coaching them on classroom observations and giving feedback to teachers. Evaluators wanted principals to know what the performance measures and indicators looked like in practice.

Districts promoted the use of professional learning communities in schools to implement school improvement strategies.

PLCs were used to align curriculum within and across grades, monitor student achievement, and create flexible student groupings for instructional purposes. In at least one district, principals required PLCs to develop their own SMART goals and monitor their success in reaching their goals.

Districts allocated or reallocated resources to help principals meet their goals.

This could be in the form of purchasing new curricular programs, hiring a coach for a principal with multiple programs in his/her school, providing resources for principals to attend workshops, ensuring there were learning specialists in the elementary schools to support instruction, or using software to efficiently observe teacher's classroom instruction and provide timely feedback to them.

PART IV.

RECOMMENDATIONS

THE FOLLOWING RECOMMENDATIONS address both challenges and issues faced by evaluators and principals during implementation of the new principal evaluation process. They are drawn from suggestions offered by the two sets of case study interviews and from the FHI 360 study team’s analysis of the two sets of survey and case study data.

The recommendations focus on two specific areas: the first on strengthening implementation of the principal evaluation process and the second on how to strengthen the design of the example model. FHI 360 believes that work in these areas will significantly strengthen the impact of principal evaluation as a positive lever for strengthening principals’ ability to improve student outcomes.

Although state and district implementation recommendations are provided separately here, most of the recommendations would require collaboration and clear communication between the state and districts to foster aligned actions that support the design and/or implementation of the model.²¹

²¹ More details on these state and district recommendations appear in Appendix A

1. Recommendations for State Actions to Strengthen Implementation

Determine a feasible strategy to accelerate the timeline for reporting state assessment results back to schools.

Credible data should be delivered back to the districts and schools by June so that summative ratings and professional growth plans can be determined before the start of the next school year. Many states struggle with delivering student assessment data in a timely manner so a feasible strategy will require a dedicated planning process.

Develop a communications strategy and products differentiated for key stakeholders to increase statewide commitment to the principal evaluation process focused on professional growth as a key means for improving student outcomes.

This could include a short report on “Why Great Principals Matter” and strategies and materials to reach important stakeholders in Minnesota.

Encourage districts to modify and streamline the example model within the parameters of the statute to match their specific context and needs.

Because the model allows some flexibility in implementation, districts should modify it within its parameters to fit their needs and contexts.

Encourage district leaders to view the next several years of early implementation as a developmental early implementation stage for refining their district’s principal evaluation process.

²² Districts should be experimenting and learning from their implementation efforts.

Encourage districts to provide data from interim student achievement measures and measures of progress on district goals to inform the goal-setting process, the mid-year conference, and the end-of-year summative conference.

Throughout the pilot study, districts increasingly turned to the use of interim student assessment data to identify progress on principals professional and school goals.

Build the capacity of districts and evaluators to implement effective principal evaluations

focused on professional growth to improve student outcomes (see also District Actions below).

Identify promising practices related to implementation of principal evaluation.

These could be collected and shared on the MDE Educator Excellence Web site as well as through statewide or regional conferences.

Engage the service cooperatives, Centers of Excellence, and state professional associations to provide training in some of the implementation strategies for key components of the example model as well as help improve the effectiveness of both principals and evaluators.

This could include sharing promising practices and direct training to strengthen model implementation, including: building evaluators’ skills in observing and providing feedback to principals, and building principals’ skills to understand what the performance measures look like in practice.

²² MDE has followed Implementation Science principles to roll out both the new teacher and new principal evaluation systems. The 2012–2013 principal evaluation pilot year was considered, in Implementation Science terms, as either an Exploration stage (for example, assessing context and fit) or an Installation stage (for example, re-allocating resources as needed) depending on the readiness or capacity of the district.

2. Recommendations for District Actions to Strengthen Implementation

Ensure that evaluators have sufficient time to make principal evaluation a priority and that they have the skills to provide feedback and learning supports to principals. Through the case studies it was evident that the success of the principal evaluation process hinges on the skills of the evaluators and their commitment to the goal of supporting principals' professional growth by building trusting relationships with them.

Ensure evaluators and principals know what the performance measures and indicators look like in practice. Districts that felt successful in implementing the rubrics for the example model offered multiple opportunities for principals to better understand the performance measures and rubrics. This could involve devoting administrative meetings to tailoring the wording of the rubric descriptors to fit the district context, regularly observing principals and providing feedback, and sharing promising practices across the district.

Encourage evaluators to talk with principals about what is working and not working in the model and make appropriate adjustments. In the case study districts, effective adaptations were made to the evaluation model because of honest conversations between evaluators and principals.

Ensure data and evidence informs the goal-setting process, the mid-year conference and the end-of-year conference. The most "data-informed" districts had access to prior year's student proficiency and growth results, which helped principals set their goals. They also had interim student assessment proficiency and growth measures to assess progress on principal's goals during the mid-year conference, as well as sufficient evidence and observations, including stakeholder survey results, to provide an accurate and fair summative rating.

Align district policies, structures, and practices to support principals' growth. This could include linking principals with key district staff to help them achieve their goals,²³ ensuring there is a data system to provide timely information and evidence, and fostering district practices on how to close the achievement gap.²⁴

Engage the support of school boards. The example model asks school boards to ensure evaluations are conducted every year.

²³ A recent Wallace Foundation report describes these actions in more detail. See *Rethinking Leadership*. October 2013. (<http://www.wallacefoundation.org/knowledge-center/school-leadership/district-policy-and-practice/Pages/Rethinking-Leadership-The-Changing-Role-of-Principal-Supervisors.aspx>)

²⁴ See, for example, The Central Office Transformation Toolkit. (New York: The Wallace Foundation.) November 2013. (<http://www.wallacefoundation.org/knowledge-center/school-leadership/district-policy-and-practice/Pages/Central-Office-Transformation-Toolkit.aspx>)

3. Recommendations for State Actions to Strengthen the Design of the Example Model

The example model provided flexibility in how they used to model to evaluate principals. Some districts implemented the example model with minimal changes, while other districts implemented other principal evaluation processes. For the most part, each principal evaluation model used by the pilot districts appeared to make inroads on achieving its primary purpose—promoting principals’ professional growth and effectiveness. Nonetheless, given that this was a pilot of the example model, both the results of the surveys and the case study interviews suggest how to strengthen the design of the example model. Many of the following ideas were suggested by principals and evaluators in the pilot districts and also appear in the recommendations above. Here, they are organized by the critical model components and pertain to actions that could be taken by MDE in collaboration with other statewide partners, stakeholders and support providers.

ORIENTATION AND TRAINING

Train new evaluators and principals in the example model. Emphasize the sharing of promising practices.

Develop a “refresher orientation” that includes revised forms and materials as well as more district- and school-level examples and promising practices. Encourage both evaluators and principals to attend these sessions, which could be offered online to those attending for a second time.

Train evaluators in effective coaching and feedback skills, possibly through the professional associations, regional cooperatives, and/or Centers of Excellence.

Encourage districts to align the performance measures and indicator rubric descriptors to meet their local needs and school improvement strategies. This might include condensing the number of indicators or focusing on a select set of performance measures each year. Districts might also identify specific activities a principal must conduct under each indicator that correlates with the district’s school improvement strategies.

Keep districts informed of updates to the MDE Educator Excellence Web site and possible changes to the example model.²⁵ This should include the posting and updating of recording forms, model materials, promising practices, training videos, and a calendar of professional learning opportunities across the state.

SELF-ASSESSMENT AND GOAL SETTING

Determine a feasible strategy for accelerating the timeline for reporting the MCA and MMR results, ideally by June of each school year.

Streamline and reformat the indicators and rubrics. Consider reducing the number of indicator descriptors and ensure each descriptor is “threaded” across all four performance levels.²⁶

Collect and share examples of promising practices of

- Self-assessment and goal-setting conferences (such as pre-conference preparation, how to organize the conference and types of feedback to provide)
- Principals’ professional and school improvement goals, including how targets were set, what measures were used, and how goal attainment was calculated (including examples of how data were used to help principals set goals)
- Performance measure rubric descriptors to help principals understand what these descriptors (and thus the indicators) look like in practice

Provide examples of interim assessments that districts might use to monitor student academic achievement and growth. (See Table 9 in this report to see the interim assessments being used in the four case study districts.)

STAKEHOLDER SURVEY DATA

Develop or commission three short stakeholder surveys (parents, teachers, students) aligned to the state performance measures and indicators, as well as to research-based school improvement strategies.

Norm these surveys with Minnesota districts.

Share examples of other short stakeholder surveys aligned to the Minnesota performance measures and indicators.

Create a handbook on administering stakeholder surveys, analyzing the results and using survey results in both goal setting and in a principal's final rating.

Gather input from districts to determine the best timing for stakeholder survey administration—before the mid-year conference to inform the principal at that stage or just prior to the end of the year so a full academic year can be considered? Or both?

MID-YEAR CONFERENCE

Provide examples of effective mid-year conference preparation and implementation (what pre-conference preparation was done, what types of evidence were brought to the meetings, how the conferences were organized, including questions asked, types of feedback that were provided by the evaluator) and how progress on goals was being assessed.

Provide guidance on using recent stakeholder survey data, if available, as evidence of progress on goals.

END-OF-YEAR SUMMATIVE CONFERENCE AND FINAL RATING

Provide examples of effective end-of-year summative conference preparation and implementation. This could include what pre-conference preparation was done, what types of evidence were brought to the meeting, how the conferences were organized, types of feedback that were provided by the evaluator) and how ratings were calculated. Include strategies districts used to address disagreements between evaluators and principals on ratings.

Provide examples from districts on how they calculated the ratings on the three components. Highlight roles principals may have played in helping to design the rating system.

Provide examples of how the Developing category can be effectively used with various principal situations. Districts in the four pilot case study sites did not use this rating and were often unaware of it.

Questions for Further Investigation

Some of the questions posed during the study remain unanswered with satisfactory certainty because of the relatively small size of the pilot survey and case study district samples. A larger-scale monitoring and feedback study could help answer the following questions and provide further insights and concrete recommendations.

I. How can the example model be used to help the professional growth of principals all along the experience continuum, from new principals to veterans?

The skill development needs of principals in relation to the principal evaluation process will vary. Although it seems clear that new principals are likely to need more guidance (a case study evaluator cautioned: “I think that [to assume that] this skill is just [going to] develop because we have the instrument is not true for the emerging principal.”), even veteran principals may need some targeted support to effectively use the new approach to principal evaluation.

II. What level of fidelity to the example model is essential for the evaluation process to produce the desired results? When do adaptations stray beyond the model’s framework?

We found that districts’ ability to make adaptations helped district’s “own” the model, strengthening their commitment to its implementation overall. However, in certain areas the acceptable amount of flexibility is unclear. For example, can a district focus on a handful of indicators each year, while still providing summative rating of principals on all the performance measures and indicators? Can a district use its interim student assessment and growth measures for the final summative rating and not wait for the MCA and MMR data to arrive in the fall? An ideal level of fidelity likely can be determined through raising and answering questions like these and then sharing across districts promising practices that are within “limits” of the model.

III. Under what conditions will both evaluators and principals perceive evaluators’ direct observations of principals’ practice as accurate and fair?

Overall, the survey responses show strong alignment between evaluator and principal perceptions of the most accurate sources of data, including evidence of progress and outcomes of their professional growth goals, direct observation of principals’ practice, self-assessment results, and mid-year conference results and feedback. The one variance was in the perception of direct observation as fair; this discrepancy in the perceived fairness of direct observations between evaluators and principals needs further study.

IV. How should the example model be translated into and supported by district policies?

The example model includes 10 district requirements (such as be consistent with a principal’s job description, incorporate district achievement goals and targets, be linked to professional development) for model implementation.²⁷ These are important and helpful requirements, but district policies should also include a district timeline for the principal evaluation process and required training of evaluators in effective coaching and feedback skills. Implications for potential district policies and the need for the state to collaborate with and support districts can be found in the recommendation sections above.

²⁷ Measuring Principal Performance in Minnesota: A Model and Recommendations. February 1, 2012, p.5.

In Summary

The pilot implementation of the new state principal evaluation process was successful in identifying and understanding:

- District approaches to principal evaluation and the adaptations that were made to the model
- How both evaluators and principals viewed the model and its components
- How participation in the model affected their practices and professional skills

FHI 360's study found many positive outcomes and potential outcomes of the use or adaptation of the example model. The study was also a valuable opportunity to gather feedback from evaluators and principals, which FHI 360's incorporated in its analysis and recommendations on how to strengthen the model implementation and design and questions about the model that need further study.

APPENDIX A.

**RECOMMENDATIONS
FOR STATE ACTION
TO STRENGTHEN
IMPLEMENTATION**

1. Recommendations for State Action to Strengthen Implementation

State Assessment Results

Identify a process for determining how to accelerate the timeline for online testing and reporting of the MCA and data analysis. Credible data should go back to the districts and schools by June so that summative ratings and professional growth plans can be determined before the start of the next school year.

Communications

Create a communication strategy and materials with input from Minnesota’s professional associations and other state education leaders to help educators and other key stakeholders understand:

- The recent research on the importance of principals’ skills in raising student achievement
- The dual intent and importance of the model to (a) promote principal’s growth and (b) ensure their accountability
- How the principal evaluation process is intended to work
- How the process has already benefitted principals and evaluators in improving their skills and practice

Experimentation, Learning, and Streamlining

Encourage district leaders to treat the first years of implementation as stages of development and implementation—and learning and adjusting from this. This can also include modifying and streamlining the model (such as rubrics and data aggregation and recording forms) for effective use in their districts.

District Capacity and Promising Practices

Identify and share promising practices related to the implementation of principal evaluation:

- Collect and share promising practices vetted by an advisory committee or another group
- Host annual or regional meetings to share these promising practices
- Encourage the service cooperatives, Centers of Excellence, and state professional associations to share these practices through their networks
- Share these practices on the MDE Educator Effectiveness Web site

Service Cooperatives, Centers of Excellence, and State Professional Associations

Collaborate with MN's professional associations, service cooperatives, and Centers of Excellence to:

- Expand the MDE Educator Excellence Web site to house and share the following:
 - a. Background information about the model including: legislative intent, relevant scholarly literature, the model, forms changes made to the model along with the rationale, the ideal timeline of model implementation
 - b. Communication strategies and materials for principals and evaluators
 - c. Approved stakeholder (parents, students, and teachers) survey tools, including a stakeholder survey administration guide
 - d. Webinars and training materials
 - e. Master state calendar of professional learning opportunities aligned with the performance measures and indicators
- Encourage Minnesota's professional associations to provide workshops or trainings for their membership based on the skills evaluators and principals need to be effective. These could include workshops on the following:
 - a. Alignment and goal setting, including:
 - » Approaches to aligning a district's strategic plan, mission, vision, and goals with schools' mission, vision, and goals and principals' goals
 - » Using data to identify and create principals' professional goals and school improvement goals; creating SMART goals and targets
 - » Aligning rubrics to a principal's performance measures and district's school improvement strategies

-
- b. Skills effective evaluators need, including:
 - » Effective strategies for coaching principals such as cognitive coaching; providing feedback to principals and building trust
 - » Using district administrative meetings to deepen principals' understanding of the evaluation process
 - » How to conduct effective observations of the principal's practice
 - » Using stakeholder survey results to help set principal's professional goals
 - » How to administer stakeholder surveys so the results are credible (sample size, response rate, etc.)
 - » How to develop a final rating process and calculation; characteristics of fair and accurate processes

 - c. Skills effective principals need, including:
 - » Strategies for using data effectively to provide formative feedback and assess principals' performance (that is, interim measures of student growth, proficiency, and gap reduction; using measures predictive of the MCA; other measures of student growth; administration, analysis, and application of the stakeholders surveys)
 - » Designing and leading effective professional learning communities (PLCs) as a mechanism to achieve school goals

2. Recommendations for District Action to Strengthen Implementation

-
- Evaluators' Time and Skills** Ensure the district evaluator(s) have the necessary time and skills to implement the model and provide coaching, supports, and feedback to the principals.
- Train evaluators in effective coaching and feedback skills and how to create trusting relationships with principals
 - Help evaluators connect principals with the resources they need to achieve their goals

-
- Performance Measures in Practice** Deepen evaluators and principals understanding of what the performance measures look like in practice by:
- Using district administrative meetings to ensure principals have a thorough understanding of the model, performance measures, indicators, and rubrics
 - Using school data to set and assess progress on principals' professional goals and school improvement goal attainment
 - Setting a schedule to observe principals in different situations and providing timely feedback to them
 - Sharing effective principals practices with other principals in the district

Data and Evidence

Establish a district data infrastructure and support for its use that provides:

- Interim student achievement measures that can provide progress monitoring during the year
- Disaggregated data to inform principals on which sub-groups in their schools may need additional support or interventions
- Stakeholder survey data aligned to the state performance standards and indicators and that are also aligned to the district's school improve strategies

Align District Policies and Supports

Align district policies, infrastructure and practices to support principals' growth by:

- Ensuring district policies and supports are in place to support the evaluation model²⁸
 - Ensuring district policies are in place to address principals who receive an unsatisfactory summative rating. (See state statutes 122A.40 and A22A.41 for guidance)
 - Developing an appropriate role for school boards to support the evaluation process such as ensuring the evaluation process occurs annually and with fidelity to the evaluation model in their district
-

²⁸ See page 9 of Measuring Principal Performance in Minnesota, February 1, 2012 for six district local decisions.

This report was written by Jane Armstrong and Deanna Burney with Maud Abeel, Constanica Warren, and Laura Dukess.

ABOUT FHI 360

FHI 360 is a nonprofit human development organization dedicated to improving lives in lasting ways by advancing integrated, locally driven solutions. Our staff includes experts in health, education, nutrition, environment, economic development, civil society, gender equality, youth, research, technology, communication and social marketing — creating a unique mix of capabilities to address today's interrelated development challenges. FHI 360 serves more than 70 countries and all U.S. states and territories.

ABOUT THE BUSH FOUNDATION

The Bush Foundation invests in great ideas and the people who power them. Established in 1953 by 3M executive Archibald Bush and his wife, Edyth, the Foundation encourages individuals and organizations to think bigger and think differently about what is possible in communities across Minnesota, North Dakota, South Dakota and the 23 Native nations that share the same geographic area.

For more information or a copy of the full report, please contact study team leads Constanica Warren or Maud Abeel at FHI 360, www.fhi360.org or (212) 243-1110.

