CAREI
UNIVERSITY OF MINNESOTA

Center for Applied Research
and Educational Improvement

# Minnesota State Teacher Development, Evaluation, and Peer Support Model Evaluation Report

*January 2015*

*Evaluation Team*
*Beverly J. Dretzke, PhD*
*Debra Ingram, PhD*
*Kristin Peterson, MA*
*Timothy Sheldon, PhD*
*Kyla Wahlstrom, PhD*

*Research Assistants:*
*Julia Baker*
*Anne Crampton*
*Elyse Farnsworth*
*Alicia Zhi Hoon Lim*
*Shannen Yap*

# TABLE OF CONTENTS

# ♦ EXECUTIVE SUMMARY

Minnesota Statutes 122A.40, Subdivision 8 and 122A.41, Subdivision 5, require that all districts evaluate teachers beginning in the 2014-2015 school year. In response to the statutes, the Minnesota Department of Education (MDE) convened a work group in early winter 2011 to consult with the MDE Commissioner to develop a state model for teacher growth and development. In winter 2013, MDE released the Minnesota State Teacher Development, Evaluation, and Peer Support Model (hereafter "Model") and began planning for a pilot of the Model during the 2013-2014 school year (hereafter "Pilot").

The Model consists of three components for evaluating teacher performance: 1) teacher practice, 2) student engagement, and 3) student learning and achievement. Sixteen school districts and one charter school from across Minnesota agreed to participate in the Pilot. Six of the participants implemented the full Model (all three components) and 11 implemented one or two components. Fourteen of the 17 districts are located outside of the Twin Cities metro area. Total enrollments ranged from 202 to 7,510 students.

In August 2013, the Joyce Foundation funded the University of Minnesota's Center for Applied Research and Educational Improvement (CAREI) in the College of Education and Human Development to conduct an evaluation of the Pilot implementation. All surveys and interview protocols utilized in the evaluation were developed by CAREI, in consultation with MDE staff. Data collection was carried out from March to June, 2014, when the Pilot year was concluding.

The Minnesota Statutes define a "teacher" as any school district staff member who is on a teacher contract. Included in this definition are individuals in traditional classroom teaching positions as well as professionals, such as librarians, nurses, school psychologists, and speech language pathologists, whose work largely takes place in non-classroom settings. Data were collected from individuals in both types of teaching positions. In this report, these positions are referred to as "classroom teacher" and "non-classroom teacher." This report summarizes responses given by classroom teachers, non-classroom teachers, summative evaluators, and MDE staff. The Executive Summary presents results of the evaluation arranged by components of the Model, implementation processes, and data source.

## ❖ Teacher Practice

The teacher practice component comprises 45% of a teacher's performance rating. Teacher practice is defined by the four domains of planning, instruction, environment, and professionalism. Teacher performance related to these domains is measured using the Minnesota Standards for Teacher Practice rubric and evidence gathered primarily from points of contact (e.g., classroom observations, one-on-one meetings) with summative evaluators and peer reviewers, the teacher's Individual Growth and Development Plan (IGDP), the teacher's self-assessment, and an optional teacher portfolio.

# Points of Contact

Points of contact are opportunities for teachers to meet with a peer reviewer or a summative evaluator to discuss professional goals, to review student learning goals, to be observed, or to evaluate their instructional practice after an observation. In the design of the Pilot Model, teachers are to meet with peer reviewers three times and with a summative evaluator two times during each school year.

## *Teacher and Summative Evaluator Surveys*

- According to the Model, a complete point of contact is grounded in the teacher's IGDP and the performance standards, is face-to-face, and is documented. The survey did not present the required elements of a point of contact, but simply asked how many points of contact the teachers had had with summative evaluators and peer reviewers during the school year. Based on their survey responses, classroom teachers had an average of 2.4 points of contact with a summative evaluator and non-classroom teachers had an average of 1.8. For points of contact with a peer reviewer, classroom teachers had an average of 3.9 while non-classroom teachers had an average of 3.7.

- Sixty percent of classroom teachers gave a rating of *good* or better to points of contact for the purpose of assigning a teacher's performance rating compared to 88% and 72%, respectively, of summative evaluators and classroom teachers who gave similar ratings. Sixty-two percent of classroom teachers gave a rating of *good* or better to points of contact for the purpose of identifying areas for professional development compared to 100% and 72%, respectively, of summative evaluators and non-classroom teachers. Therefore, for both purposes, classroom teachers' ratings of points of contact were less positive than those of summative evaluators and non-classroom teachers.

## *Teacher Interviews*

- The majority of teacher interviewees (70%) mentioned benefits of increased collaboration and stronger relationships among faculty resulting from points of contact with peer reviewers.

- Those peer reviewers who did classroom observations voiced the challenge of leaving their own students even if a substitute teacher was offered for that time period. Although they commented on how much they learned from each other when they were able to visit one another's classrooms, they also requested clarification regarding their role. Was the function of the peer reviewer that of a mentor or an evaluator?

- While the points of contact with peer reviewers were perceived as useful, teachers found them terribly time-consuming when it came to documentation. The paperwork was thought to be unreasonable and repetitive.

- Another challenging aspect of points of contact with peer reviewers reported by some teachers was that they did not feel comfortable being honest with their peers about what they

saw in the observations. In situations where peer reviewers were reluctant to give honest feedback, collegial relationships did not flourish despite increased contact between teachers.

- Teachers talked about varying levels of quality in the points of contact with summative evaluators. Key factors included the amount of time allotted to the pre- and post-observation conferences and the specificity of the feedback. Teachers appreciated when the points of contact were focused and specific, and several attributed this to the use of the IGDP and the rubric.

*Summative Evaluator Interviews*

- The points of contact were seen by summative evaluators as a useful tool because they validated professional conversations, and they shifted the evaluation system away from being only a compliance and regulatory system to one of professional growth.

- Summative evaluators' negative perceptions of points of contact were primarily centered on the paperwork needed to document their occurrence. The comment most frequently expressed by the summative evaluators was their desire to do more classroom observations than observations of PLC or grade-level meetings because classroom observations provided more meaningful information.

## Individual Growth and Development Plan (IGDP)

*Teacher and Summative Evaluator Implementation Surveys*

- Classroom teachers, in general, spent more time preparing their IGDPs than did non-classroom teachers. The average time spent by classroom teachers was 2.4 hours compared to an average of 1.9 hours spent by non-classroom teachers. Summative evaluators reported spending an average of 55 minutes per teacher reviewing IGDPs.

- Teachers were encouraged to incorporate district goals, school goals, or professional learning community goals into their IGDPs. More than three fourths of classroom teachers and non-classroom teachers reported that they referred to one or more of these goals during the development of their IGDPs.

- As part of their IGDPs, teachers identify at least one professional growth goal and the evidence that will be used to evaluate goal achievement. Approximately half of classroom teachers and non-classroom teachers indicated that they encountered difficulties when writing their professional growth goals. The writing of performance standards, however, was perceived to be more difficult for classroom teachers than for non-classroom teachers, with 60% and 41%, respectively, indicating that they found the task to be difficult.

- In preparing their goals, teachers were more likely to consult with a peer reviewer than with a summative evaluator. Approximately one third of classroom teachers and non-classroom teachers reported that they consulted with a summative evaluator compared to 77% of classroom teachers and 83% of non-classroom teachers who reported that they consulted with

a peer reviewer. Both types of consultation were considered beneficial by the majority of teachers.

- The vast majority of both classroom (76%) and non-classroom teachers (84%) indicated that the development of professional growth goals helped them to reflect on their instructional practice.

- Survey participants were asked to rate the information generated by the IGDP for the purpose of assigning a valid rating for a teacher's performance. Ratings given to the IGDP by summative evaluators and non-classroom teachers were higher than those given by classroom teachers. More specifically, 53% of classroom teachers gave a rating of *good* or better to the use of the IGDP for this purpose compared to 75% and 73%, respectively, of summative evaluators and non-classroom teachers.

- Survey participants were also asked to rate the information generated by the IGDP for the purpose of identifying areas where a teacher might benefit from professional development. Again, summative evaluators and non-classroom teachers gave higher ratings than classroom teachers, with 54% classroom teachers giving a rating of *good* or better to the use of the IGDP for this purpose compared to 72% and 82%, respectively, of summative evaluators and non-classroom teachers.

### *Teacher Interviews*

- The majority (75%) of teachers found the IGDP to be very useful, with the remaining teachers expressing more neutral feelings about the process. The most common reason expressed about why the IGDP worked well for teachers was that it kept them focused. Reasons for less enthusiastic responses included dissatisfaction with their professional goal, such as setting a narrow goal that was too easily achieved. Teachers also found the IGDP less useful if they felt swamped by other components of the Model.

- Overall, teachers found the self-assessment process useful, although some thought it was overly time-consuming.

### *Summative Evaluator Interviews*

- Summative evaluators referred to the value of having a structure in the form of the IGDP to facilitate conversations with teachers. They explained that the IGDP enabled the teachers to choose their own areas for personal and professional growth, and thus allowed principals to engage in meaningful, reflective conversations with each of them about their work.

# ❖ Student Learning Goals (SLG)

The student learning and achievement component comprises 35% of a teacher's performance rating. This component is made up of student learning goals (SLGs) and data from a value-added assessment. The value-added assessment is not included in this evaluation report since value-added results will not be available until February 2015.

MDE defines an SLG as a measurable, long-term student academic growth target that the teacher sets at the beginning of the year. SLGs are intended to demonstrate a teacher's impact on student learning within a specified interval of instruction. Several elements must appear in a teacher's SLG, including the standards the goal will align with, the assessments to be used to measure goal attainment, the time period covered by the goal, and expected student growth. All teachers (classroom and non-classroom) were required to set learning goals, select an assessment, establish mastery scores, and establish student start points.

## *Teacher and Summative Evaluator Implementation Surveys*

- On average, classroom teachers spent 2.5 hours working on their SLGs while non-classroom teachers spent 1.7 hours. Summative evaluators reported that they spent an average of 50 minutes per teacher working on SLGs.

- Overall, 42% of teachers indicated that they found the setting of SLGs to be an easy task. Of the tasks related to setting SLGs, the establishment of student start points was easier for the teachers than selecting an assessment or establishing mastery scores.

- Teachers were twice as likely to consult with a peer reviewer (69%) than with a summative evaluator (33%) in the development of their SLGs. Regardless of the individuals consulted, the majority of teachers reported that the consultations were beneficial.

- Summative evaluators tended to view SLGs more positively than classroom teachers did with respect to SLGs generating information to be used to assign a rating to a teacher's performance or to identify areas where a teacher might benefit from professional development. Approximately three fourths of summative evaluators gave positive ratings to the use of SLGs for these two purposes compared to approximately 56% of classroom teachers. The opinions of non-classroom teachers were somewhat more positive than those of classroom teachers.

## *Teacher Interviews*

- Over 45% of the participating teachers used an internal measure to set their class learning goal, whereas just over one third used an external measure such as the Minnesota Comprehensive Assessments (MCA) or the Benchmark Assessment System (BAS). Most of the remaining teachers (15%) also used internal measures, but they sought "improvement" rather than specific quantified results.

- Forty percent (40%) of the classroom teachers interviewed said SLGs were useful. The teachers who found SLGs useful talked about the way the goals increased their focus, planning, and reflection.

- Specialist teachers (i.e., art, physical education, family and consumer science, music, etc.) found the SLGs to be somewhat challenging because of complications due to subject matter, larger numbers of students, and fewer meetings with all students in any given week.

- Nearly all special education teachers interviewed found the development of SLGs to be redundant, given the fact their students already have individual learning goals on their Individualized Educational Plans (IEPs).

- Roughly half of the teachers interviewed felt their students had met the class learning goal they had established, although only 18% had supporting data. Those without supporting data who felt the class learning goal had been achieved spoke about being satisfied with the progress of their students.

- In addition to setting SLGs for the entire classroom, teachers were also expected to set a goal for a smaller, targeted group of students. Two thirds of the teachers interviewed said they had set targeted group goals, while one third did not designate a targeted group. Twenty percent of teachers talked about the fact they did not have an adequate number of students to warrant both class learning goals and targeted learning goals. Teachers in most districts were neutral or unenthused about drawing distinctions between groups of students.

- Teachers who commented on problems with the SLGs (45% of those interviewed) found them unwieldy for a number of reasons, particularly noting a need for more guidance and training on how to measure growth and then use that data to change instructional practice.

*Summative Evaluator Interviews*

- The implementation of SLGs was seen by all summative evaluators as being only moderately successful at best.

- The positive feelings about SLGs concerned the setting of measureable learning targets for students. Principals thought the data from the learning goals provided concrete information and a good conversational focus for where teachers could improve.

- The negative perceptions were centered not on the concept of setting learning targets, but on the need for more training for some teachers on developing and setting measureable goals based on standards.

- Issues related to how specialists and non-classroom teachers (e.g., librarians, nurses) struggled with setting measureable student learning goals were mentioned by all interviewed summative evaluators.

# ❖ Student Engagement Component

The Model stipulates that results of a reliable and valid student survey will make up 15% of the student engagement evidence and an additional 5% will come from other measures of engagement (e.g., observations, videos) for a total of 20% of a teacher's performance rating. If a survey is not available, as would likely be the case for teachers of lower primary grades, then the entire 20% will be made up of other measures of engagement. Two versions of a State Model Student Survey were developed to assist districts with the assessment of student engagement. One version was for students in grades 3 to 6 and the other for students in grades 6 to 12.

## *Teacher and Summative Evaluator Implementation Surveys*

- The two most common methods utilized by classroom teachers to assess student engagement were student surveys—the State Model Student Survey or a different survey (100%) and observations (89%).

- The average time spent on assessment of student engagement by classroom teachers and non-classroom teachers was 2.9 hours and 1.6 hours, respectively.

- The summative evaluators reported that they spent an average of 2.4 hours on activities related to the State Model Student Survey.

- Teachers were asked their opinions about using student engagement evidence in their performance evaluations and their responses indicated that they were not enthusiastically supportive. More specifically, 57% of teachers gave a rating of *poor* or *fair* to the use of student engagement assessments for assigning a rating to a teacher's performance. Fifty-seven percent also gave a rating of *poor* or *fair* to the use of student engagement assessments for identifying areas where a teacher could benefit from professional development.

## *Teacher Beliefs Survey*

A teacher beliefs survey was administered to classroom teachers at the beginning of the school year before the Pilot was administered in their districts and again at end of the school year after student survey feedback reports had been received. The purpose of the surveys was to address questions concerning teachers' opinions about the use of student surveys before implementation of the Pilot and how their opinions had changed after student survey results had been received and reviewed with the summative evaluators.

- At the beginning of the school year, most teachers indicated that they perceived student feedback to be potentially valuable for providing information that would help them improve their effectiveness. More specifically, 55% of teachers said they were in favor of using student surveys to evaluate their teaching, and 78% said they believed that student survey feedback would enable them to improve their teaching effectiveness.

- On the survey administered at the beginning of the school year, teachers indicated they were uncertain about their students' ability to provide valid feedback. The most common concerns

expressed by teachers on the fall survey were that students would use the survey to get back at teachers and that students would give low ratings in subjects that were difficult for them.

- Near the end of the school year, after student survey feedback reports had been received, the teachers were significantly less supportive of including student surveys in their evaluations and significantly more concerned about the validity and usefulness of student survey feedback. Although their support had decreased, however, the teachers had not become strongly opposed to student surveys, but rather more skeptical. On the spring survey, 45% of teachers indicated they were in favor of using student survey feedback to evaluate their teaching and 67% indicated they felt student surveys would provide them with information that would enable them to improve their effectiveness.

- When compared to teachers, the principals seemed to place more value on feedback from student surveys, both for the purpose of assigning a rating to a teacher's performance and for identifying areas where a teacher could benefit from professional development.

*Teacher Interviews*

- Of the teachers interviewed, 40% were positive about the student survey experience, and a few more (45%) thought a similar survey should be used in the future. Positive responses tended to touch upon the importance of hearing students' voices.

- Forty-two percent of teachers expressed concerns about validity of the survey data, citing the age of their students as potentially affecting their ability to comprehend the survey items or value the survey's purpose.

- An additional area of concern was the way the State Model Student Survey data compared teachers against one another. This seemed unnecessary to many of the teachers, and at least one teacher at every interview district expressed concern for their colleagues in the event of low student engagement results.

*Summative Evaluator Interviews*

- The student engagement component had the widest range of opinions about its usefulness of any of the components in the Model, with 70% of summative evaluators expressing mixed views. The positive reactions centered on the value of including the student's voice in assessing a teacher's performance.

- Negative comments, specifically about the student engagement survey, centered around two main themes—difficulty in scheduling the administration of the survey and concerns about the wording of questions, especially for the younger students.

# ❖ MDE-Provided Resources

*Teacher and Summative Evaluator Implementation Surveys*

- Overall, opinions about the adequacy of MDE training were more positive than negative, with most teachers indicating that they thought that MDE training provided adequate guidance for developing their IGDPs and their SLGs. However, teachers indicated that training would have been more effective if an overview of the whole process was provided before step-by-step instructions were given on individual components. They also indicated they would have benefited from training that included more time on examples that fulfilled the requirements.

- Summative evaluators were equally divided between being neutral and satisfied regarding the training and support received for the State Model Student Survey administration and interpretation. Thus, overall, it appears that the summative evaluators found the training and support to be adequate.

- The majority of classroom (62%) and non-classroom teachers (58%) perceived the Minnesota Standards for Teacher Practice rubric to be useful for identifying areas of their practice for professional growth. The majority of classroom (58%) and non-classroom teachers (67%) also perceived the rubric to be useful for identifying areas of their practice for evaluation.

*Teacher Interviews*

- Perceptions about the Minnesota Standards for Teacher Practice were decidedly split between those who found it useful and those who did not. Those who spoke positively mentioned that it served as a structure on which to build and grow and that it would provide stability over time. Those who spoke negatively believed that the rubric limited thinking and led to canned ways of reflecting about their work.

*Summative Evaluator Interviews*

- The summative evaluators viewed the Minnesota Standards for Teacher Practice rubric as a useful feature of the evaluation process that should be continued because it has an emphasis on growth and support. The few concerns that were expressed mostly related to vague language in some of the standards and the lack of time and training for them, as leaders, to effectively use the rubric to stimulate professional growth conversations.

- Summative evaluators said that they received good training from MDE with respect to the State Model Student Survey. However, several principals commented they would have benefited from more training specifically focused on helping teachers interpret the student survey results.

# ❖ Perceptions of the Model Overall

### *Teacher and Summative Evaluator Implementation Surveys*

- The benefit of the Model most frequently identified by both teachers and summative evaluators was that the Model facilitated professional conversations, specifically, teacher-to-teacher conversations and summative evaluator-to-teacher conversations about instructional practice.

- Summative evaluators' ratings of the Model overall were significantly more positive than those given by either classroom teachers or non-classroom teachers. The teachers' lower ratings appeared to be related mostly to the time required to implement the Model and aspects of the Model that teachers found to be confusing or cumbersome.

- Teachers and summative evaluators were in strong agreement regarding the changes that should be made to the Model to make it more effective. Their top three recommendations, in order of highest to lowest, were: a) streamline the Model; it is too time-consuming, b) simplify the Model; it is too complex, and c) clarify the Model; it is too difficult to follow the instructions or see how the parts fit together.

- With respect to areas needing improvement, teachers pointed out the infrequency and limited scope of points of contact, the absence of an introductory overview of the entire process, and the lack of comprehensive follow-through on the part of some summative evaluators.

### *Teacher Interviews*

- Overall, 59% of the teachers mentioned that a benefit of the Model was the focus on their instructional practice. Many teachers indicated that the Model asked them to do things they were already doing on their own, but with more structure and follow-through. However, 75% mentioned the amount of time it required as a challenge. In addition, a number of teachers were very concerned about what would happen in the future when their district would not have financial support from the state to cover implementation costs.

### *Summative Evaluator Interviews*

- Nearly all summative evaluators described positive perceptions of changes in their school's culture as a result of participating in the Pilot. Two main themes emerged from their comments. One theme was that a sense of professionalism among the teachers was created by having school-wide participation in the teacher evaluation Pilot. The other was that the teacher evaluation system was leading to greater accountability among the teachers.

- All summative evaluators commented that they had learned an enormous amount by participating in the year-long Pilot, despite the overwhelming quantity of required paperwork. Furthermore, most summative evaluators stated that they were planning to take many of the features of the Model and use them in the creation of a modified version of a teacher evaluation system for their district's use in the next school year.

# ❖ MDE Staff Interviews

Individual interviews with two Minnesota Department of Education staff responsible for the development and implementation of the Pilot were conducted in July and August 2014. The titles of the MDE staff members are: Educator Effectiveness Supervisor and the Director of the Division of School Support. The responses from the two interviewees are combined in the summary presented below.

## *Successful Aspects of the Pilot Model Implementation*

The MDE interviewees described success as occurring on two levels. First, the overall Model appears to have effectively encouraged teacher buy-in to an evaluation process, and the second was learning which aspects of the Model were most successfully implemented and why. The willingness of the Pilot districts to participate and to cooperate with MDE was identified as essential for the Pilot's valuable outcomes.

Additionally, the Pilot districts were effective at implementing components of the Model where summative evaluators and peer reviewers shared feedback with teachers and offered guidance on using the feedback to improve instructional practice. This process was perceived to have increased teacher buy-in to the evaluation Model. Points of contact and feedback given to teachers were viewed as the most successful elements of the evaluation system.

## *Areas for Future Development of the Model*

The implementation of the Pilot was a learning experience that helped identify aspects of the Model where modifications would be beneficial. An idea that surfaced multiple times was having a phased implementation strategy, rather than rolling out all pieces of the Model at once. A phased strategy might reduce confusion and feelings of being overwhelmed by the Model's expectations.

Possible revisions to specific aspects of the Model were also identified. For example, it was mentioned by MDE staff that the number of points of contact might be reduced. It was also suggested that the peer review process be analyzed to identify the training and ongoing coaching that teachers need in order to become effective as peer reviewers. The peer review process was considered to be one of the most valuable elements of the Model, and improved peer reviewer training could make it even more so.

One surprising outcome was the difficulty that teachers had with the creation and measurement of student learning goals. It was suggested that increasing training on how to set, assess, and reflect on data from student learning goals would likely eliminate much of the difficulty.

A common theme in summative evaluator feedback received by MDE staff was that implementation of the Model's multiple elements was a significant time commitment. When asked about this, both MDE interviewees noted that processes generally take longer when an initiative is first implemented than later on after it has become more routine. After the Model becomes familiar and summative evaluators and teachers have acquired experience with it, it is

likely that the components will take less time and effort. It is also likely that summative evaluators will identify ways that the time commitment can be reduced without affecting the quality of the evaluation process

Several lessons were learned from the Pilot of the teacher assessment model. It is expected that evaluation models will continually evolve, incorporating technology in new ways and shifting to meet the ever-changing needs and goals of the educational system. It is expected that evaluation models will be tailored to specific schools and will be adjusted as roles and responsibilities become better defined.

To support the process of Model refinement, a focus on the big picture and deep understanding of the purpose should be emphasized in training and implementation. It is also important to continue to offer training and professional development to ensure continued fidelity of implementation and a strong emphasis on the overarching purposes of the Model.

A district's evaluation model should be integrated with overall goals and existing initiatives in order for the district to benefit fully from the time and effort put forth by school administration and staff. A meaningful balance across all activities must be reached to prevent the participants from becoming overwhelmed.

Overall, the Model requires a great deal of communication within and across groups and roles. Establishing new routines and expanded communication channels will take time, practice, and a belief that these will lead to increases in student performance. Changes in teacher practice will necessitate continuing discussions among teachers and principals to ensure that the changes result in improved education for all students.

# Minnesota State Teacher Development, Evaluation, and Peer Support Model Evaluation Report

## ♦ OVERVIEW

Minnesota Statutes 122A.40, Subdivision 8 and 122A.41, Subdivision 5, require that all districts evaluate teachers beginning in the 2014-2015 school year. In response to the statutes, the Minnesota Department of Education (MDE) convened a work group in early winter 2011 to consult with the MDE Commissioner to develop a state model for teacher growth and development. In winter 2013, MDE released the Minnesota State Teacher Development, Evaluation, and Peer Support Model (hereafter "Model") and began planning for a pilot of the Model during the 2013-2014 school year (hereafter "Pilot").

The Model consists of three components for evaluating teacher performance: 1) teacher practice, 2) student engagement, and 3) student learning and achievement. Sixteen school districts and one charter school from across Minnesota agreed to participate in the Pilot. Six of the participants implemented the full Model (all three components) and 11 implemented one or two components. Fourteen of the 17 are located outside of the Twin Cities metro area. Total enrollments ranged from 202 to 7,510 students.

In August 2013, the Joyce Foundation funded the University of Minnesota's Center for Applied Research and Educational Improvement (CAREI) in the College of Education and Human Development to conduct an evaluation of the Pilot. All surveys and interview protocols utilized in the evaluation were developed by CAREI, in consultation with MDE staff.

The Minnesota Statutes define a "teacher" as any school district staff member who is on a teacher contract. Included in this definition are individuals in traditional classroom teaching positions as well as professionals, such as librarians, nurses, and speech language pathologists, whose work largely takes place in non-classroom settings. Data were collected from individuals in both types of teaching positions. In this report, these positions are referred to as "classroom teacher" and "non-classroom teacher." This report summarizes responses given by classroom teachers, non-classroom teachers, summative evaluators, and MDE staff. The Executive Summary presents results of the evaluation arranged by components of the Model, implementation processes, and data source.

# ♦ TEACHER AND SUMMATIVE EVALUATOR IMPLEMENTATION SURVEYS

## Methodology

### *Teacher Implementation Survey*

#### *Instrument*

The questionnaire contained questions designed to collect teachers' perspectives on the usefulness of the Model's elements, sources of support they had received, difficulties encountered, time spent on specific tasks, and their overall opinion of the Model. Items were grouped into the following sections: Individual Growth and Development Plan, Student Engagement, Student Learning Goals, Overall Evaluation, and Background Information. The background information items included the grade level(s) taught during the 2014-15 school year, subject area, and total number of years teaching. Various closed-option response scales were used to collect teacher responses. Items requesting open-ended comments were also included. The survey instrument can be found in Appendix C.

#### *Procedure*

MDE provided a list of district contacts who were then emailed and asked to provide a list of teacher emails. If the district contact did not respond, teacher email addresses were found on the MDE or district website. A message describing the survey and its purpose was emailed to the teachers, along with a link to the on-line survey. The initial emails were sent to teachers on May 28, 2014. A reminder email was sent to non-respondents on June 12. The survey closed at the end of the day on Friday, June 20.

#### *Participants*

Invitations were emailed to a total of 866 teachers, and 353 completed the Teacher Implementation Survey, resulting in a response rate of 41%. The sample was comprised of 335 classroom teachers and 18 staff members who were on teacher contracts but did not hold a traditional classroom teaching position. Examples of these staff are librarians, nurses, and speech language pathologists. Throughout the report, the respondents who were not traditional classroom teachers are referred to as "non-classroom teachers."  Additional descriptive information regarding the respondents is provided in Appendix D.

## Summative Evaluator Implementation Survey

### Instrument

The Summative Evaluator Implementation Survey was designed to gather information about the summative evaluators' activities related to the Model and their perceptions regarding the Model's components. The questionnaire items were grouped into the following sections: Information About Your Building, Teacher Practice, Student Learning Goals, Student Engagement Survey, Communication and Training, and Satisfaction with the Model. Various closed-option response scales were used to collect summative evaluator responses. In addition, open-ended items provided the opportunity for summative evaluators to write comments or suggestions they had regarding the Model. The survey instrument can be found in Appendix E.

### Procedure

An email invitation was sent on May 27, 2014, to all 40 individuals in the Pilot districts who were serving as summative evaluators. The message contained a description of the research study and a link to the survey. Responses were collected between May 27 and June 16. The first reminder email was sent to non-responders on June 2, and a final reminder was sent on June 10. The survey closed at the end of the day on Friday, June 20.

### Participants

Thirty of the 40 summative evaluators responded to the survey, yielding a response rate of 75%. The respondents represented 11 unique districts. Fifty percent of the respondents reported that they were the only summative evaluator in their building, 41% reported that their building had two summative evaluators, and the remainder reported there were more than two. The number of teachers in the summative evaluators' buildings ranged from 6 to 140 (M = 33.86, SD = 28.56). The number of teachers for whom the respondents were serving as summative evaluator ranged from 3 to 46 (M = 20.82, SD = 13.45).

## Findings of the Teacher and Summative Evaluator Implementation Surveys

This section of the report summarizes the survey responses of classroom teachers, non-classroom teachers, and summative evaluators. The findings related to teacher practice are presented first, followed by student learning goals, and then student engagement.

## *Teacher Practice*

Results regarding the teacher practice component are presented separately for each of the component's three elements in the following order: a) Individual Growth and Development Plan (IGDP), b) points of contact, and c) performance standards for the teacher practice rubric.

### *Individual Growth and Development Plan (IGDP)*

The IGDP is intended to be an organized way for a teacher to set and pursue professional growth goals and to plan related learning activities as part of a three-year professional review cycle. The teachers' professional growth goals should be aligned with and support district, school, and professional learning community goals. Teachers were expected to consult with their assigned peer reviewer during the development and revisions of their IGDPs, and the assigned summative evaluator must annually approve the Plan and its revisions.

Individuals who responded to the teacher implementation survey were asked to provide information regarding the time they spent completing their IGDP, resources and consultants they utilized, the difficulty of completing tasks related to the IGDP, and their perceptions of the validity and usefulness of the evidence provided by the IGDP with respect to enabling a summative evaluator to rate a teacher's performance and identify areas where a teacher might benefit from professional development. In this report, the IGDP is also referred to as the "Plan."

*Time spent preparing the Plan*. All non-classroom teachers and 99% of classroom teachers reported that they had completed their IGDP. The time that the teachers spent preparing their Plans is summarized in Table 1. In general, the classroom teachers spent more time working on their plans than the non-classroom teachers did. The average hours spent were 2.43 for classroom teachers and 1.93 for non-classroom teachers. The range of hours spent was also considerably more for classroom teachers as compared to non-classroom teachers, 0.50 to 10.00 hours versus 1.00 to 4.00 hours, respectively. Summative evaluators reported spending an average of 55 minutes per teacher reviewing IGDPs.

Table 1. Time spent preparing the IGDP (hours)

| Descriptive statistic | Classroom teacher (*n = 169*) | Non-classroom teacher (*n = 11*) |
|---|---|---|
| Mean | 2.43 | 1.93 |
| Standard deviation | 1.75 | 1.01 |
| Median | 2.00 | 1.75 |
| Minimum | 0.50 | 1.00 |
| Maximum | 10.00 | 4.00 |

***Goals referred to during Plan development***. The teachers were asked if they referred to district, school, or professional learning community goals while preparing their Plans. Responses, summarized in Table 2, indicate that more than three fourths of the teachers referred to each of these goals when creating their Plans. The percentages were somewhat higher for non-classroom teachers than for classroom teachers.

Table 2. Percent of teachers referring to district goals, school goals, and professional learning community goals during Plan development

| Type of goals | Classroom teachers | | Non-classroom teachers | |
|---|---|---|---|---|
| | *n* | *%* | *n* | *%* |
| District goals | 171 | 77% | 12 | 83% |
| School goals | 173 | 81% | 12 | 92% |
| Professional learning community (PLC) goals | 174 | 81% | 12 | 100% |

***Persons consulted during Plan development***. Only a few classroom teachers (18%) and non-classroom teachers (8%) indicated that they wrote their Plans on their own without consulting anyone. Of those who did consult with someone, both classroom teachers and non-classroom teachers were more likely to consult with a peer reviewer than with a summative evaluator (see Table 3). More specifically, 77% of classroom teachers and 83% of non-classroom teachers reported that they consulted with a peer reviewer compared to 36% and 33%, respectively, reporting that they consulted with a summative evaluator. To a lesser extent, both groups conferred with other individuals. For example, among classroom teachers, 10% consulted with another teacher or colleague and 8% consulted with a professional learning community.

Table 3. Persons consulted during Plan development

| Person consulted | Classroom teachers (*n* = 177) | Non-classroom teachers (*n* = 12) |
|---|---|---|
| | *%* | *%* |
| Peer reviewer | 77% | 83% |
| Summative evaluator | 36% | 33% |
| No one; I wrote them on my own | 18% | 8% |
| Another teacher or colleague | 10% | 17% |
| Professional learning community | 8% | 0% |

*Consultations beneficial to Plan development*. The majority of classroom and non-classroom teachers found that consultations regarding development of their Plans were beneficial whether they consulted with a peer reviewer, a summative evaluator, or someone else (see Table 4). However, it appears that consultations with peer reviewers were perceived to be especially valuable, with 91% of classroom teachers and 100% of non-classroom teachers stating that peer reviewer consultations were beneficial compared with 64% and 75%, respectively, indicating that summative evaluator consultations were beneficial.

Table 4. Percent of classroom teachers and non-classroom teachers indicating that consultations were beneficial

| Person consulted | Classroom teachers | | Non-classroom teachers | |
|---|---|---|---|---|
| | *n* | % | *n* | % |
| Peer reviewer | 144 | 91% | 10 | 100% |
| Summative evaluator | 86 | 64% | 4 | 75% |
| Other | 55 | 76% | a | a |

ªOnly one non-classroom teacher responded to the item.

*Difficulty of writing professional growth goals*. Teachers were asked to rate the level of difficulty they experienced when writing their professional growth goals. As shown in Figure 1, more than half the classroom teachers (56%) found that writing their professional growth goals was *somewhat* or *very difficult* while about 16% found the task to be *somewhat* or *very easy*. In contrast, half the non-classroom teachers perceived the task to be *somewhat difficult* and 41% found the task to be *somewhat* or *very easy*. Based on a comparison of the response distributions, it appears that that non-classroom teachers found writing their professional growth goals to be less difficult than did classroom teachers.
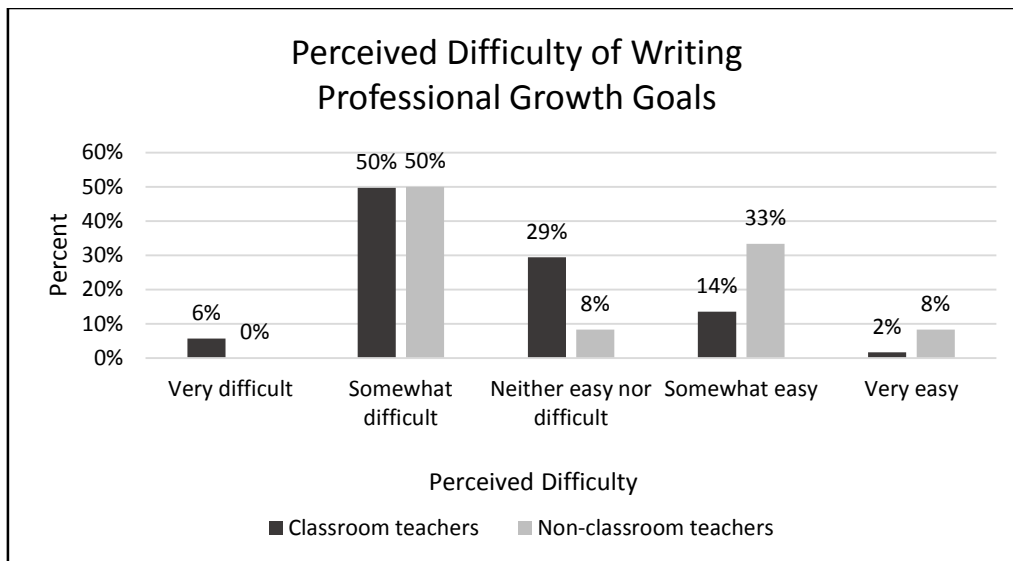


Figure 1. Classroom and non-classroom teachers' perceived difficulty of writing professional growth goals.

***Difficulty of writing performance standards for goals***. When asked about the difficulty of writing performance standards for their goals, 60% of classroom teachers reported that the task was *somewhat* or *very difficult*, and 18% reported that the task was *somewhat* or *very easy* (see Figure 2). The distribution of responses given by non-classroom teachers was quite different, with 41% of non-classroom teachers reporting the task to be *somewhat* or *very difficult*, and 25% reporting the task to be *somewhat easy*. Similar to responses regarding difficulty of writing professional growth goals, classroom teachers tended to find writing performance standards to be more difficult than did their non-classroom counterparts.
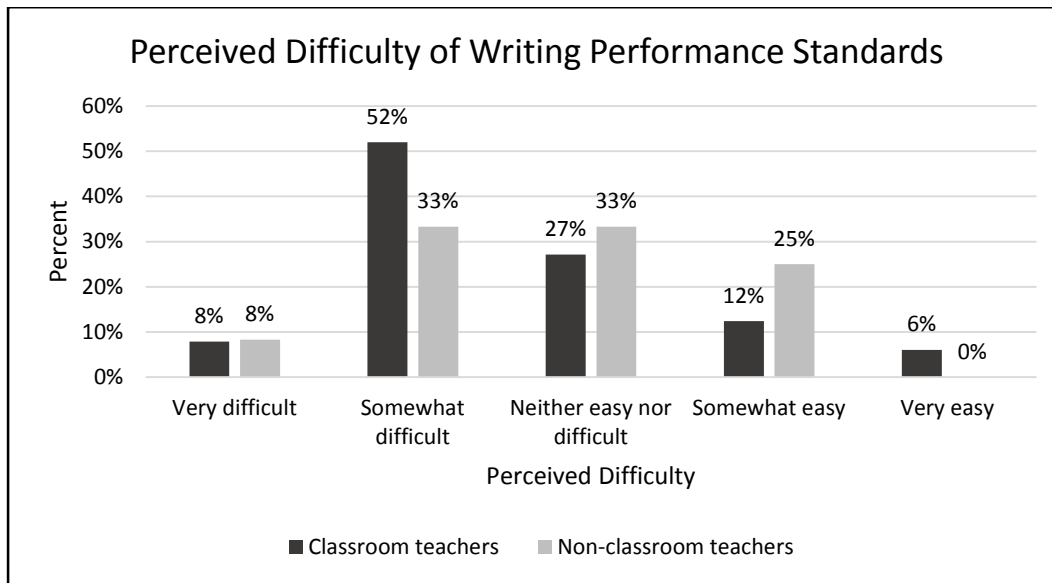


Figure 2. Classroom and non-classroom teachers' perceived difficulty of writing performance standards.

   ***Teacher Development, Evaluation, and Peer Support Model: Implementation Handbook***. The teachers were asked whether MDE's Implementation Handbook provided adequate guidance for developing their Plans. Although the most frequent response given by both classroom (36%) and non-classroom teachers (60%) was *somewhat true*, non-classroom teachers seemed to view the Handbook more positively than did classroom teachers (see Figure 3). More specifically, 41% of classroom teachers agreed that the Handbook provided adequate guidance compared to 60% of non-classroom teachers.
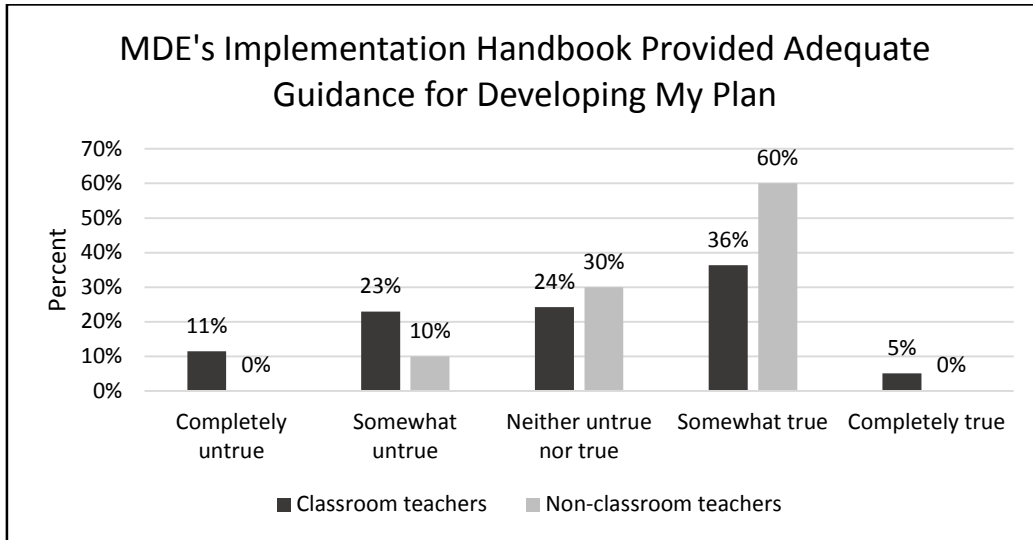


Figure 3. Classroom and non-classroom teachers' perceived utility of the Handbook for providing guidance for developing their IGDP.

   ***Training provided by MDE***. MDE staff provided on-site training to all but one of the Pilot districts. That district opted to offer their own training. The results regarding the teachers' opinions of the MDE training, therefore, do not include responses from teachers in that district. Because excluding that district's data left very few non-classroom teachers in the sample, only the responses of classroom teachers are summarized in Figure 4. Overall, opinions about the adequacy of MDE training were more positive than negative, with 48% of classroom teachers indicating that they thought that MDE training provided adequate guidance for developing their Plans.
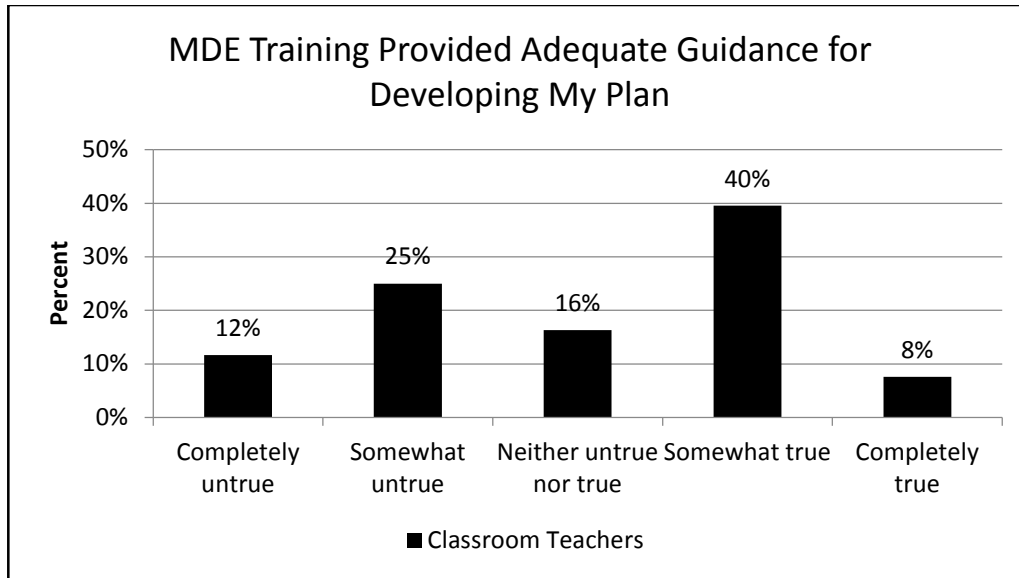
Figure 4. Classroom and non-classroom teachers' perceived utility of the MDE training for providing guidance for developing their Plans.

*Professional growth goals and reflection on instructional practice*. Respondents were asked whether developing professional growth goals helped them to reflect on their instruction practice. The results were very favorable (see Figure 5). Seventy-six percent and 84%, respectively, of classroom teachers and non-classroom teachers indicated that developing their professional growth goals did help them to reflect on their instructional practice.
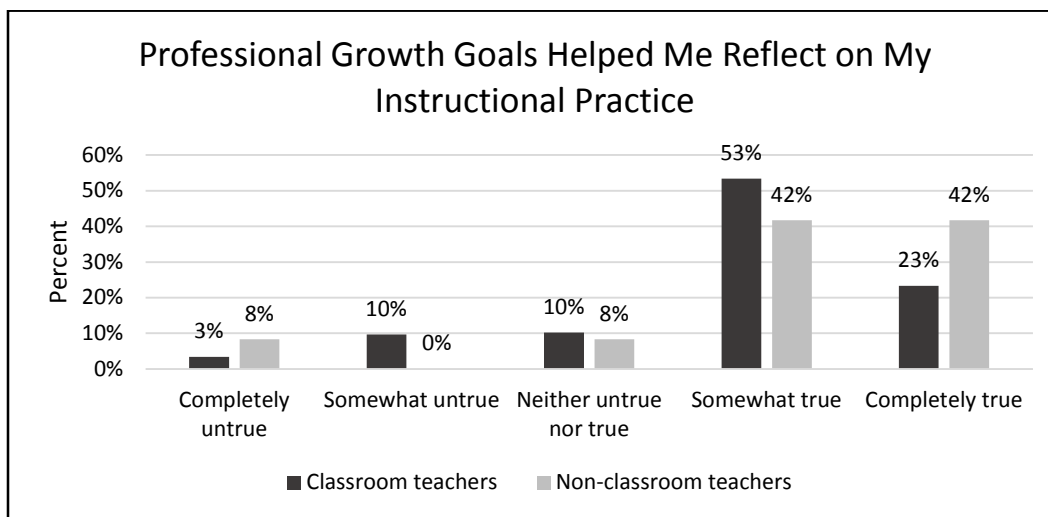


Figure 5. Classroom and non-classroom teachers' responses regarding whether professional growth goals helped them to reflect on their instructional practice.

   ***Rating the IGPD for generating information that enables the summative evaluator to assign a valid performance rating***. Teachers and summative evaluators were asked to rate the IGDP with respect to generating information that enables the summative evaluator to assign a valid rating to a teacher's performance. The response scale ranged from *poor* to *excellent*. Overall, the ratings were quite positive with more than half of classroom teachers, non-classroom teachers, and summative evaluators selecting a rating of *good*, *very good*, or *excellent*. However, a comparison of the distributions in Figure 6 shows that non-classroom teachers and summative evaluators gave higher ratings than the classroom teachers. More specifically, 73% of non-classroom teachers and 75% of summative evaluators gave a rating of *good, very good,* or *excellent* compared to 53% of classroom teachers.
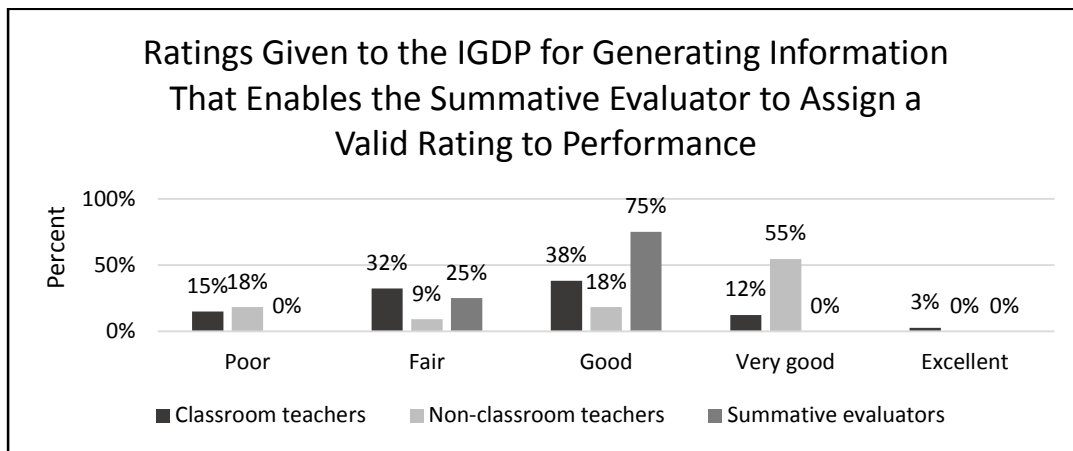


Figure 6. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of the IGDP for generating information that enables the summative evaluator to assign a valid rating to a teacher's performance.

*Rating the IGPD for generating information that enables the summative evaluator to identify areas for professional development*. Respondents were also asked to rate the IGDP with respect to generating information that enables the summative evaluator to identify areas where a teacher might benefit from professional development. The majority of classroom teachers, non-classroom teachers, and summative evaluators indicated that they believed the IGDP would provide useful information regarding professional development needs. Similar to the results displayed in Figure 6, a comparison of the distributions displayed in Figure 7 shows that the ratings given by non-classroom teachers and summative evaluators were generally more positive than the ratings given by classroom teachers. More specifically, 82% of non-classroom teachers and 72% of summative evaluators selected a rating of *good*, *very good*, or *excellent* compared to 54% of classroom teachers.
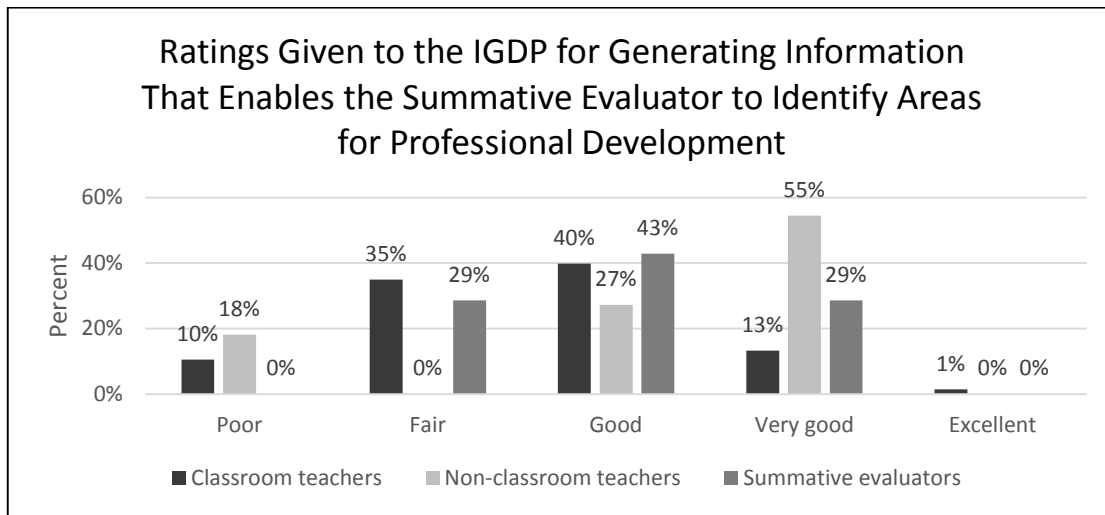


Figure 7. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of the IGDP for generating information that enables the summative evaluator to identify areas where a teacher might benefit from professional development.

*Open-ended comments*. Survey respondents were given the opportunity to write open-ended comments about the IGDP. Comments were provided by 37 teachers (classroom and non-classroom combined) and 3 summative evaluators. A summary of teachers' open-ended comments is presented first, followed by a summary of the summative evaluators' comments.

● *Teachers' open-ended comments*. Comments written by teachers were analyzed by topic, and the topics that occurred more than twice are summarized here. These topics are:  a) role of the summative evaluator, b) effectiveness of the training, c) time required to carry out tasks, d) ways in which the process was confusing, and e) knowledge gained from experience with the Model. Topics are summarized in order of highest to lowest frequency of occurrence.

*Role of the summative evaluator (12 comments)*. Comments regarding the role of the summative evaluator mentioned what the summative evaluator did not do, limitations of the IGDP for the summative evaluator, and the usefulness of the IGDP for the summative evaluator. Example comments for each of these subtopics are presented below.

### What the summative did not do:
- *My summative evaluator never rated my plans.*
- *My summative evaluator didn't do anything with my goals because he wasn't required to. That is a direct quote.*

### Limitations of the IGDP for the summative evaluator:
- *The observer sees just small piece of what teacher is doing in regards to their goals.*
- *I do not believe my growth plan allows the summative evaluators to know anything about my strengths or weaknesses in the classroom. My growth plan is for me to measure my own successes and failures. If I am being honest with myself when filling it out and choose a goal I truly struggle with, what they see is where I struggle. I may choose a goal that focuses on something that I'm already quite good at, and I will be great. In each case what the summative evaluator will see is quite different. To make things more difficult, some goals have little to do with classroom practices and I had trouble finding ways to get the evaluators to see what I doing in those areas. It's a great idea, but easily manipulated.*

### Usefulness of the IGDP for the summative evaluator:
- *This provided the specific guidelines/areas for the evaluator to observe.*
- *It would be good as long as my summative evaluator was flexible and allowed me to monitor and adjust my plan as the understanding of student and teacher needs developed.*

**Effectiveness of the training (8 comments)**. Comments about the effectiveness of training primarily concerned the Handbook, desirability of an overview of the whole process, and requests for specific training experiences. Example comments for each of these subtopics are presented below.

### The Handbook:
- *The handbook was very confusing and overwhelming.*
- *The manual is very thick and not user friendly.*

### Desirability of an overview of the whole process:
- *We had to write our goals before we were given training on what the overall picture would look like for this process and our goal did not fit well in the end.*
- *It seemed like each time we were told about a new type of goal. It seemed had we known about each of them, we could have lined them up together a lot better.*

### Requests for specific training experiences:
- *During training, too much language used to explain the process and why doing it. Not enough time spent on just generating examples of goals.*
- *I would like to have more professional development on 1. writing a properly sized plan (one that is meaningful and doable) and 2. on strategies for collecting and tracking the student data I need to measure the effectiveness of my plan.*

**Time required to carry out tasks (7 comments)**. Teachers' comments about time either expressed that tasks took a lot of time to complete or there was too much paperwork. Example comments are shown below.

*Tasks took a lot of time:*
- *This plan takes a lot of time to prepare. Couldn't be done all in one sitting. Had to go back and rethink and evaluate.*
- *This whole process was EXTREMELY difficult and very time consuming. My peer teacher was extremely helpful, but this job made her life pure hell with overload of work.*

*Too much paperwork:*
- *In general, there was too much redundant paperwork.*
- *I think all of this could be done without involving so much time and paperwork.*

**Ways in which the process was confusing (6 comments)**. In their comments, the teachers described various ways in which the process was confusing. The first comment shown below describes confusion related to one goal turning into four goals, and the second describes confusion related to an individual's position as a pre-school teacher.

- *I found the goal, which turned into 4 goals, to be very confusing. It was not made clear with the first meeting we had with the state team.*
- *I consulted with my high school leader and she was able to help a little but the high school is doing something that is a whole different age level* [than pre-school]*, so I am still on my own... I am also still unsure where I fit for the school wide and district goals.*

**Knowledge gained from experience with the Model (3 comments)**. The three teachers who commented on knowledge gained from experience with the Model indicated they now have the knowledge to complete the process more effectively or, in hindsight, would have written their Plans differently.

- *Knowing what I know now, I would have written my IGDP differently.*

● *Summative evaluators' open-ended comments*. Only three summative evaluators wrote comments regarding the IGDP. One of these individuals pointed out that most of the plans he/she reviewed did not have measurable goals. Another individual stated that participating in the Pilot was a learning experience that will lead to an improvement in procedures. The third individual described how tasks related to teachers' Plans were distributed between two summative evaluators.

*Points of Contact*

Points of contact are opportunities for teachers to meet with a peer reviewer or a summative evaluator to discuss professional goals, to review student learning goals, to be observed, or to evaluate their instructional practice after an observation. Points of contact with peer reviewers

are intended to be private exchanges between colleagues. The summative evaluator is not to be present, and information gathered during these exchanges is not to be shared with the summative evaluator. On the other hand, points of contact with summative evaluators are private exchanges that become part of the teacher's personnel record.

*Number of points of contact*. Teachers were asked how many points of contact they participated in with peer reviewers and with summative evaluators during the school year. Their responses are summarized in Table 5. For number of points of contacts with a peer reviewer, classroom teachers' responses ranged from 0 to 15 with an average of 3.93. The number of peer reviewer contacts reported by non-classroom teachers was somewhat less, ranging from 2 to 5 with an average of 3.67. As would be expected, the number of contacts with summative evaluators was less than the number of contacts with peer reviewers for both classroom teachers and non-classroom teachers. The number of summative evaluator contacts reported by classroom teachers ranged from 0 to 9 with an average of 2.41, and the number reported by non-classroom teachers ranged from 0 to 4 with an average of 1.75.

Table 5. Descriptive statistics for teachers' points of contact with peer reviewers and summative evaluators

| Type of teacher | Descriptive Statistic | Point of contact with: | |
| --- | --- | --- | --- |
| | | Peer reviewer | Summative evaluator |
| Classroom teacher | *n* of respondents | 172 | 173 |
| | Mean | 3.93 | 2.41 |
| | SD | 2.05 | 1.84 |
| | Mode | 3 | 3 |
| | Minimum | 0 | 0 |
| | Maximum | 15 | 9 |
| Non-classroom teacher | *n* of respondents | 12 | 12 |
| | Mean | 3.67 | 1.75 |
| | SD | 1.16 | 1.29 |
| | Mode | N/A | 2 |
| | Minimum | 2 | 0 |
| | Maximum | 5 | 4 |

*Rating points of contact for generating information that enables the summative evaluator to assign a valid performance rating*. Teachers and summative evaluators were asked to rate points of contact with respect to generating information that enables the summative evaluator to assign a valid rating to a teacher's performance. Overall, the ratings were quite favorable, with the majority in each of the three groups selecting a rating of *good* or better (see Figure 8). The summative evaluators' ratings were especially high with 88% selecting a rating of *good* or better compared to 60% and 72%, respectively, of classroom teachers selecting similar ratings.
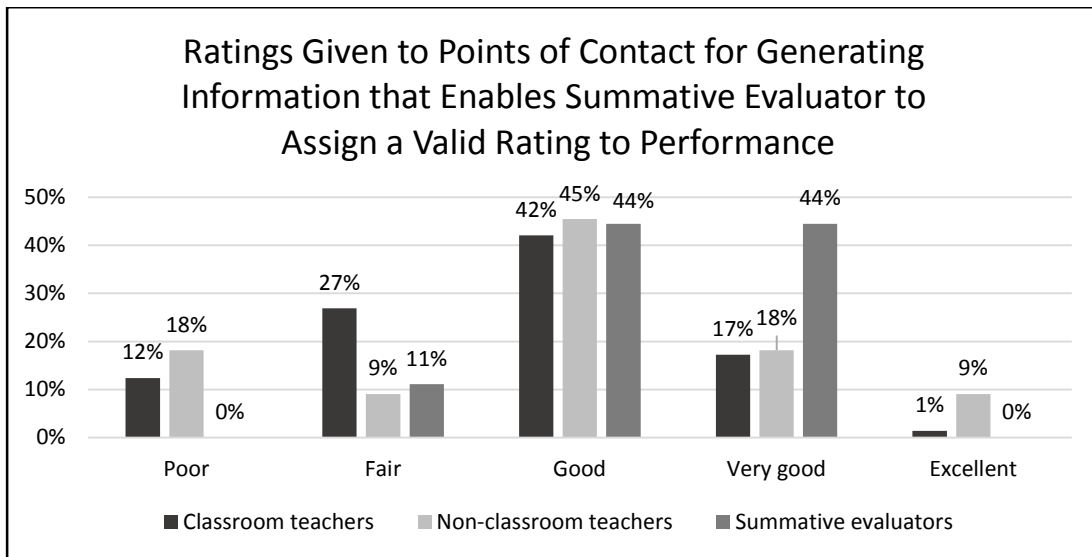


Figure 8. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of points of contact for generating information useful for assigning a valid performance rating to a teacher.

*Rating points of contact for generating information that enables the summative evaluator to identify areas for professional development*. Based on their ratings, it appears that the majority of classroom teachers, non-classroom teachers, and summative evaluators believe that points of contact are useful for enabling the summative evaluator to identify areas where a teacher might benefit from professional development (see Figure 9). In addition, it appears that summative evaluators especially feel that the information obtained from points of contact is valuable for this purpose since 100% of summative evaluators selected a rating of *good* or better compared to 62% of classroom teachers and 72% of non-classroom teachers who selected similar ratings.
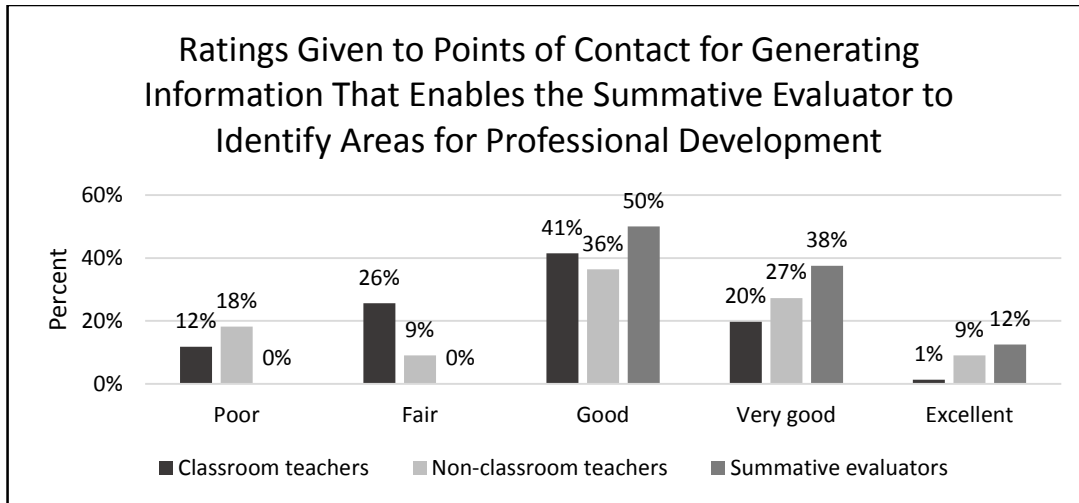
Figure 9. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of points of contact for generating information that enables the summative evaluator to identify areas for a teacher's professional development.

***Open-ended comments***. Open-ended comments about points of contacts were provided by 23 teachers and 1 summative evaluator. A summary of teachers' open-ended comments is presented first, followed by a summary of the summative evaluator's comment.

● ***Teachers' open-ended comments***. Teachers' comments were analyzed by topic, and topics that occurred more than twice are included in the summary. These topics, in order of frequency of occurrence, are: a) Limitations of points of contact, b) positive reactions to points of contact, c) confusion regarding the points of contact process, and d) dissatisfaction with the summative evaluator's behavior. One or more example comments are provided for each topic.

***Limitations of points of contacts (6 comments)***. The limitations of points of contact mentioned by teachers mainly concerned their infrequency and narrow scope or not having the time to use the information to improve one's practice. For example:

- *Times met with summative evaluator were good, but they are only one small time of our whole year and I don't believe that what my evaluator saw was my normal everyday teaching. They just saw one tiny slice of the pie that is my teaching time.*
- *Had VERY little time to actually use information obtained from the points of contact to change my practice.*

***Positive reactions to points of contact (4 comments)***. In their positive comments, the teachers stated that peer reviewers and/or summative evaluators were generally helpful or that face-to-face contacts were especially valuable.  For example:

- *Both my peer reviewer and my summative evaluator gave sufficient goals and advice.*
- *I did enjoy sitting down and talking with my administrator face to face few times a year and going over valid general info.*

*Confusion regarding the points of contact process (4 comments)*. Four teachers indicated they were confused about points of contact requirements. In addition, some teachers indicated they were not familiar with the role of the summative evaluator. For example:

- *It was unclear to us whether points of contact had to be classroom observations or not.*
- *Not sure what this all means with the peer review and the summative evaluator. Our team met weekly to review our goals so that was about 30 sessions. Our principal met with us weekly as well. So is she the summative????*

*Dissatisfaction with summative evaluator behavior (3 comments)*. Three teachers indicated they were dissatisfied with some aspect of their summative evaluator's behavior. Either they had had no points of contact with their summative evaluator or the points of contact were rushed and disjointed. For example:

- *Have not met with my summative evaluator at all this year.*
- *Many of the points of contact I had with the summative evaluators were rushed and spaced out. They would observe, wait a couple weeks, then come back to talk about it. They didn't remember the lesson and the notes they took were incomplete or vague. They also failed to do the full observation rotation (pre, during, and post), affecting their view of lesson and making it seem incomplete. It felt like they walked into the middle of the movie and then critiqued it. They can see where I struggled, but had a tough time pointing out what I had done well. It's easy for me to see my flaws, I need to know what I did right too.*

● *Summative evaluator's open-ended comments*. One summative evaluator provided a comment on the use of points of contact. This individual indicated it was difficult to be an effective summative evaluator given the numerous other responsibilities of the administrator's position. The comment is shown below.

- *To really be effective, we would need to wear less hats in our school. Two administrators doing everything makes it tough to put the time needed but we are becoming more proficient.*

### Performance Standards for Teacher Practice Rubric

The *Performance Standards for Teacher Practice Rubric* was developed by MDE to provide guidance to teachers in formulating their performance standards and goals. Survey respondents were asked if the rubric was an effective tool for a) identifying areas of their practice for professional growth and, b) identifying areas of their practice for evaluation. The response scale ranged from *completely untrue* to *completely true*.

*Identifying areas of my practice for professional growth*. Figure 10 shows that the majority of classroom teachers and non-classroom teachers perceived the rubric to be useful for

identifying areas of their practice for professional growth. More specifically, 62% of classroom teachers and 58% of non-classroom teachers indicated the rubric was useful for this purpose.



Figure 10. Classroom teachers' and non-classroom teachers' perceptions of the rubric as a useful tool for identifying areas for professional growth.

***Identifying areas of my practice for evaluation***. The majority of classroom teachers and non-classroom teachers, 58% and 67%, respectively, perceived the rubric to be useful for identifying areas of their practice for evaluation (see Figure 11). The classroom teachers' responses, however, were more variable than the classroom teachers' ratings and were spread across all five rating options whereas the non-classroom teachers' responses were only distributed among the middle three options.
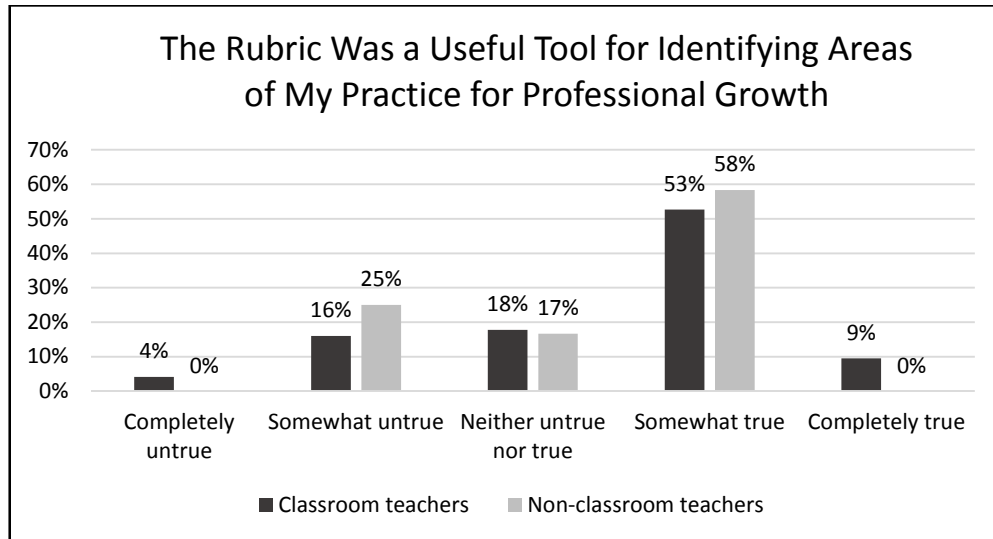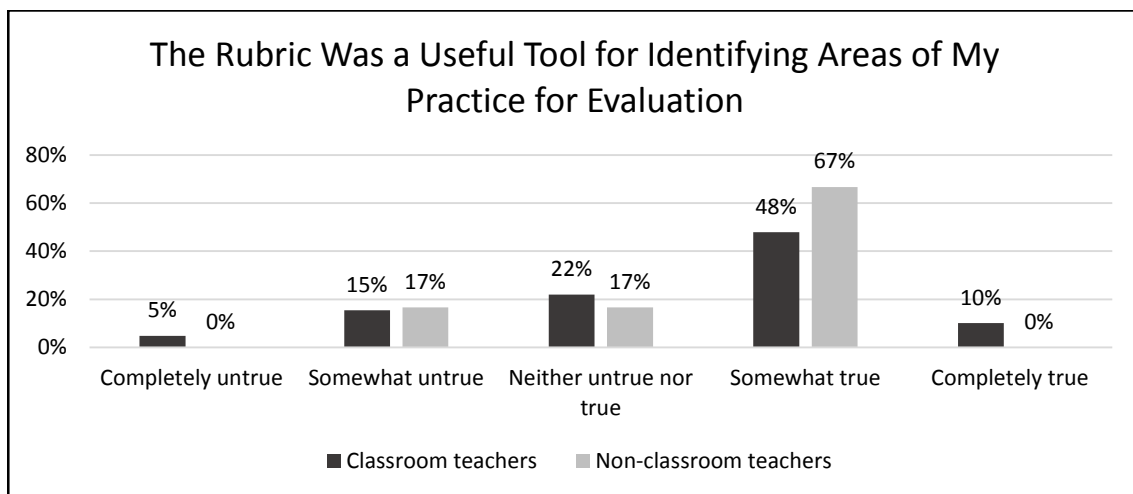


Figure 11. Classroom teachers' and non-classroom teachers' perceptions of the rubric as a useful tool for identifying areas for evaluation.

*Open-ended comments*. The teacher survey provided an opportunity for respondents to write open-ended comments about the teacher practice rubric. The comments, made by 11 respondents, were analyzed according to topic. Two topics occurred at least twice.  They were: a) Limitations of the teacher practice rubric, and b) unfamiliarity with the teacher practice rubric.

*__Limitations of the teacher practice rubric (5 comments)__*. Generally, in comments that mentioned limitations, respondents said that the rubric did not provide a good fit for their teaching situations. For example:

- *Rubric did not translate across all areas of teaching, i.e. Special Education.*
- *I think the framework was helpful but am not convinced it is the best framework to use for us as a school. There are elements of teacher practice and responsibilities not included or not tailored to our work at* [name of school].

In addition, one respondent mentioned a lack of accountability and a lack of training to help teachers reach their goals.

- *I did not take the goals I made seriously; mainly I just jumped through the hoops needed. There is no real accountability for them, and even the person who evaluated me said she would just give all good results just because it is not our job to critique performance of another teacher. At this point and time, I can recall my areas of improvement, but I did nothing with them. There was no training to HELP me get better at those goals.*

*__Unfamiliarity with the teacher practice rubric (4 comments)__*. Four teachers indicated they were not familiar with the teacher practice rubric. From their responses, it is not certain whether they did not use the rubric or if they referred to it by another name. For example,

- *Did we use that?*
- *I'm not sure what document you are referring to.*

### Summary of Teacher Practice Findings

- Classroom teachers, in general, spent more time preparing their IGDPs than did non-classroom teachers. The average time spent by classroom teachers was 2.4 hours compared to an average of 1.9 hours by non-classroom teachers. Summative evaluators reported spending an average of 55 minutes per teacher reviewing IGDPs.

- More than three fourths of classroom teachers and non-classroom teachers referred to district goals, school goals, or professional learning community goals during the development of their IGDPs.

- Approximately half of classroom teachers and non-classroom teachers indicated that they encountered difficulties when writing of their professional growth goals. The writing of performance standards, however, was perceived to be more difficult for classroom teachers

than non-classroom teachers, with 60% and 41%, respectively, indicating that they found the task to be difficult.

- In preparing their goals, teachers were more likely to consult with a peer reviewer than with a summative evaluator. Approximately one third of classroom teachers and non-classroom teachers consulted with a summative evaluator compared to 77% of classroom teachers and 83% of non-classroom teachers who consulted with a peer reviewer. Both types of consultation were considered beneficial by the majority of teachers.

- Overall, opinions about the adequacy of MDE training were more positive than negative, with 48% of classroom teachers indicating that they thought that MDE training provided adequate guidance for developing their Plans. Non-classroom teachers seemed to view the Handbook more positively than did classroom teachers. Sixty percent of non-classroom teachers and 41% of classroom teachers said that the Handbook provided adequate guidance for developing their Plans.

- The vast majority of both classroom (76%) and non-classroom teachers (84%) indicated that the development of professional growth goals helped them to reflect on their instructional practice.

- On average, during the school year, classroom teachers had 3.9 points of contact with peer reviewers while non-classroom teachers had 3.7.  The number of points of contact with a summative evaluator were somewhat less. On average, classroom teachers reported an average of 2.4 points of contact with a summative evaluator and non-classroom teachers reported an average of 1.8.

- Summative evaluators and non-classroom teachers gave higher ratings than classroom teachers to the use of information generated by the IGDP for the purpose of assigning a rating to a teacher's performance. More specifically 53% of classroom teachers gave a rating of *good* or better to the use of the IGDP for this purpose compared to 75% and 73%, respectively, of summative evaluators and non-classroom teachers.

- Summative evaluators and non-classroom teachers also gave higher ratings than classroom teachers to the use of information generated by the IGDP for the purpose of identifying areas where a teacher might benefit from professional development. More specifically, 54% classroom teachers gave a rating of *good* or better to the use of the IGDP for this purpose compared to 72% and 82%, respectively, of summative evaluators and non-classroom teachers.

- Classroom teachers gave somewhat more positive ratings to points of contact than to the IGDP as a source of information for assigning teacher performance ratings and identifying areas for professional development, but overall, their ratings were still lower than those given by summative evaluators and non-classroom teachers. More specifically, 60% of classroom teachers gave a rating of *good* or better to points of contact for the purpose of assigning a teacher's performance rating compared to 88% and 72%, respectively, of summative evaluators and classroom teachers; and 62% of classroom teachers gave a rating of *good* or

better to points of contact for the purpose of identifying areas for professional development compared to 100% and 72%, respectively, of summative evaluators and non-classroom teachers.

- Teachers identified strengths of the teacher practice component related to consultations with peers and summative evaluators, reflections on their teaching practice, and the specific guidance they received for carrying out points of contact and developing the IGDP. On the other hand, they also identified areas needing improvement related to the time taken to complete the requirements, the infrequency and limited scope of points of contact, the absence of an introductory overview of the entire process, and the lack of comprehensive follow-through on the part some summative evaluators.

- At the end of the Pilot year, the majority of classroom, non-classroom teachers, and summative evaluators approved of the teacher practice component of the Model. However, based on the survey results, their approval would be described as more reserved than whole-hearted.

## *Student Learning Goals (SLG)*

Student learning goals (SLGs) comprise one part of the student learning and achievement component of the Model. MDE defines an SLG as a measurable, long-term student academic growth target that the teacher sets at the beginning of the year. SLGs are intended to demonstrate a teacher's impact on student learning within a specified interval of instruction. Several elements must appear in a teacher's SLG including the standards the goal will align with, the assessments to be used to measure goal attainment, the time period covered by the goal, and expected student growth.

This section of the report presents results of survey items concerning SLGs. Survey questions asked respondents about the resources provided by MDE, time spent on SLGs, difficulty of carrying out tasks related to SLG formulation, consultations regarding SLGs, and the validity and usefulness of SLGs for assigning teacher performance ratings and identifying professional development needs.

### *Resources Provided by MDE to Support Teachers in the Development of Student Learning Goals*

To support teachers in developing SLGs, MDE provided the *Student Learning Goals Handbook* and training on SLG development. Teachers were asked about the adequacy of these two MDE-provided resources.

   ***Student Learning Goals Handbook***. Nearly half (47%) of classroom teachers and more than half (63%) of non-classroom teachers felt that the Handbook provided adequate guidance in SLG development (see Figure 12).
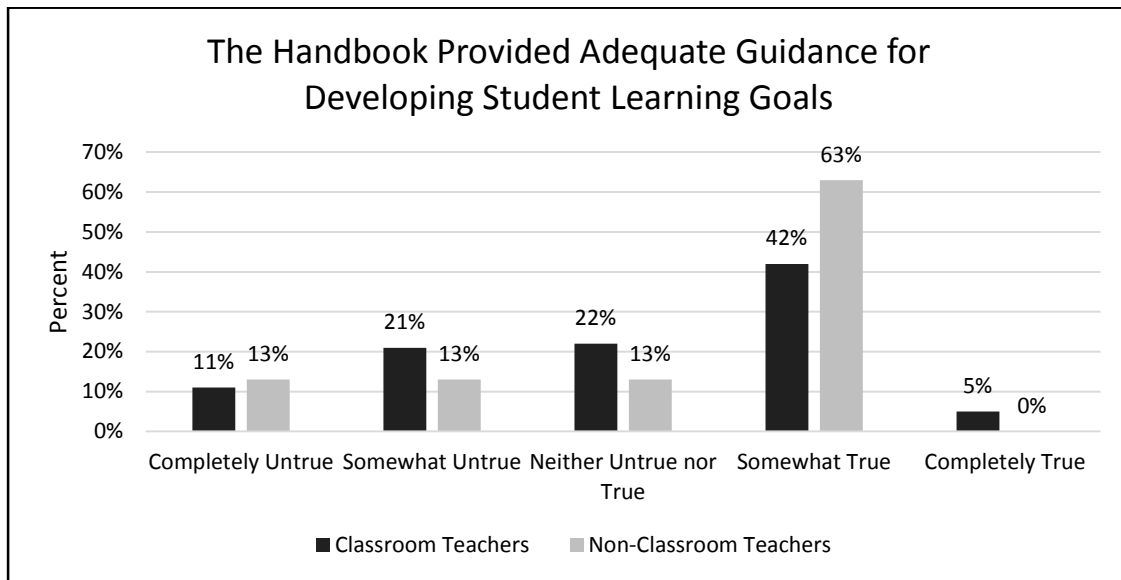


Figure 12. Responses of classroom and non-classroom teachers to the statement "The Handbook provided adequate guidance for developing student learning goals."

***Training provided by MDE***. Classroom teachers' responses regarding MDE-provided training indicated that the majority (58%) felt that MDE-provided training gave adequate guidance for developing their SLGs (see Figure 13). Note that these results do not include responses of teachers in the district that did not participate in MDE-provided training but, instead, opted to provide its own training. Also, the small number non-classroom teachers ($n = 4$) responding to the survey item did not permit presentation of the non-classroom teacher results.
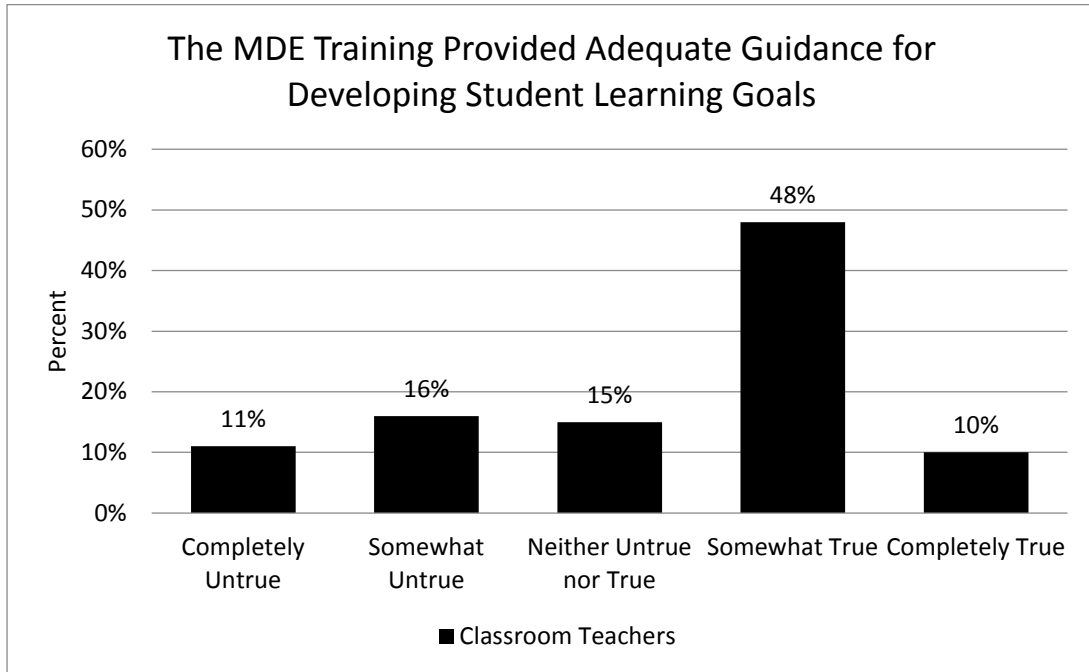


Figure 13. Classroom and non-classroom teacher responses to the item "The MDE training provided adequate guidance for developing student learning goals."

### *Time Spent on SLGs*

On average, classroom teachers reported spending more time than non-classroom teachers working on SLGs, 2.54 hours compared to 1.69 hours, respectively (see Table 6). Summative evaluators reported that they spent an average of 50 minutes per teacher. The total time spent by summative evaluators was calculated by multiplying the number of teachers supervised by the average time per teacher. The total time spent on SLGs by summative evaluators ranged from 1.50 to 40.00 hours, with a mean of 17.08 hours.

Table 6. Total hours spent working on SLGs

| Descriptive statistic | Classroom teacher *(n =183)* | Non-classroom teacher *(n = 9)* | Summative evaluator *(n = 10)* |
|---|---|---|---|
| Mean | 2.54 | 1.69 | 17.08 |
| Standard deviation | 1.76 | 0.83 | 14.58 |
| Median | 2.00 | 1.50 | 14.25 |
| Minimum | 0.00 | 1.00 | 1.50 |
| Maximum | 8.67 | 3.00 | 40.00 |

### *Perceived Difficulty of Setting SLGs*

The steps for setting SLGs are, in order, a) selecting an appropriate assessment tool, b) establishing mastery scores, and c) establishing student start points. Teachers were asked about the perceived difficulty of accomplishing these three steps and the perceived overall difficulty of setting SLGS. Response options were *very difficult*, *somewhat difficult*, *neither easy nor difficult*, *somewhat easy*, and *very easy*. Responses given by classroom teachers are summarized in Table 7. Overall, 42% of classroom teachers found the task of setting SLGs to be easy. By examining responses regarding the difficulty of accomplishing the three tasks required for setting SLGs, it is apparent that teachers were more likely to encounter difficulties with selecting an appropriate assessment and establishing mastery scores than with establishing student start points.

Table 7. Classroom teachers' perceived difficulty of tasks related to setting SLGs

| SLG task | Classroom teachers | | |
|---|---|---|---|
| | *n* | % selecting *somewhat difficult* or *very difficult* | % selecting *somewhat easy* or *very easy* |
| Setting student learning goals | 171 | 30% | 42% |
| • Selecting or developing an assessment to measure the goals | 168 | 44% | 29% |
| • Establishing mastery scores for the assessment | 166 | 51% | 28% |
| • Establishing student start points | 170 | 32% | 41% |

*Persons Consulted When Setting SLGs*

The majority of classroom teachers consulted with others in the development of their SLGs. Only 18% reported that they wrote them on their own without consulting with anyone (see Table 8). The most frequently reported consultations were with peer reviewers and summative evaluators, with peer reviewers consulted nearly two times as often as summative evaluators, 69% versus 33%, respectively. Less frequently, teachers reported consulting with other teachers or colleagues or with a professional learning community.

Table 8. Persons consulted by classroom teachers in the development of SLGs

| Person consulted | % Classroom teachers (*n* = 184) |
|---|---|
| Peer reviewer | 69% |
| Summative evaluator | 33% |
| No one; I wrote them on my own | 18% |
| Another teacher or colleague | 9% |
| Other, not specified | 7% |
| Professional learning community | 5% |

*Consultations Beneficial to Setting SLGs*

The majority of teachers found consultations regarding SLGs to be beneficial (see Table 9). However, consultations with peer reviewers appear to have been especially beneficial since 89% indicated that consultations with peer reviewers were beneficial compared to about two thirds stating that consultations with summative evaluators or other individuals were beneficial.

Table 9. Percent agreement with the item: "It was beneficial to consult with a ___ in formulating my student learning goals"

| Consultant | *n* | Percent agreeing it was beneficial |
|---|---|---|
| Peer reviewer | 142 | 89% |
| Summative evaluator | 92 | 67% |
| Other | 45 | 69% |

***Ratings of SLGs for Generating Information That Enables the Summative Evaluator to Assign a Valid Performance Rating***

Teachers and summative evaluators were asked to rate SLGs with respect to generating information that enables the summative evaluator to assign a valid rating to a teacher's performance. Using a response scale of *poor* to *excellent*, over half of classroom teachers, non-classroom teachers, and summative evaluators selected a rating of *good* or better. A comparison of the distributions displayed in Figure 19 shows that the summative evaluators' ratings were especially positive. Overall, 75% of summative evaluators selected a rating of *good* or better compared to 57% and 66%, respectively, of classroom teachers and non-classroom teachers who selected similar ratings.
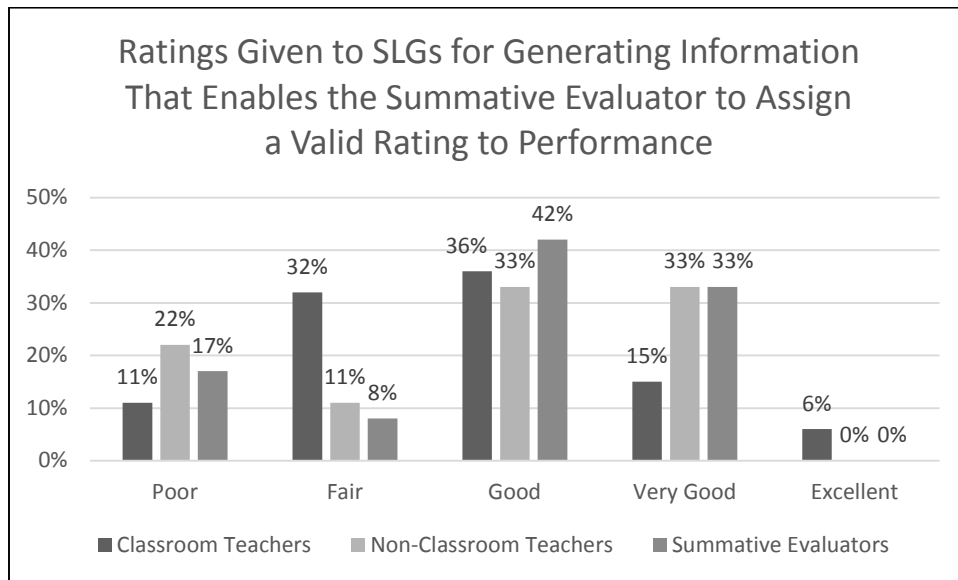


Figure 19. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of SLGs for generating information that enables the summative evaluator to assign a valid rating to a teachers' performance.

***Ratings of SLGs for Generating Information That Enables the Summative Evaluator to Identify Areas for Professional Development***.

Teachers and summative evaluators were also asked to rate SLGs with respect to generating information that enables the summative evaluator to identify areas where a teacher might benefit from professional development. Overall, the ratings were positive with the majority of classroom teachers (56%), non-classroom teachers (77%), and summative evaluators (77%) selecting a rating of *good* or better (see Figure 20).
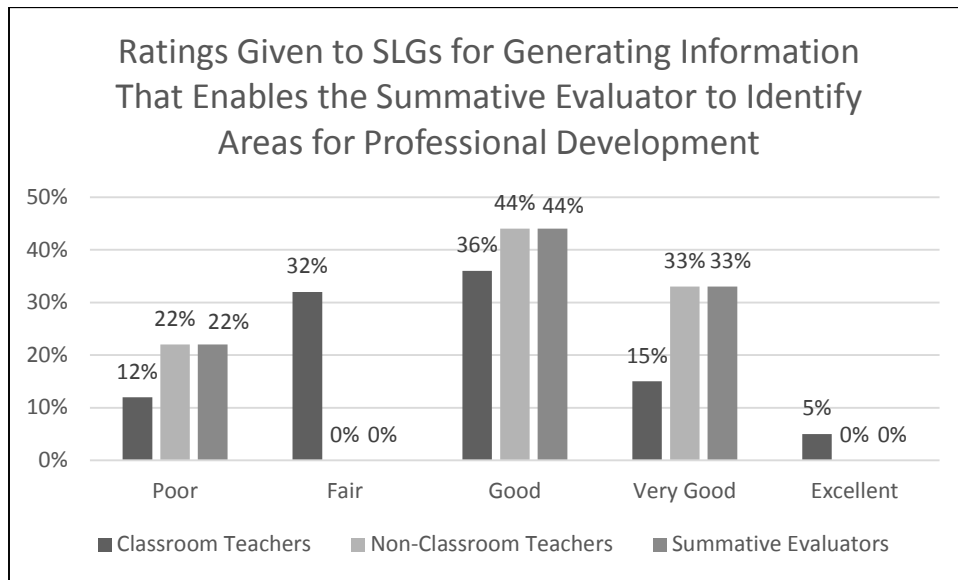


Figure 20. Classroom teachers', non-classroom teachers', and summative evaluators' ratings of SLGs for generating information that enables the summative evaluator to identify areas where a teacher might benefit from professional development.

### *Open-Ended Comments*

On the summative evaluator survey, respondents were invited to write open-ended comments regarding SLGs. Only three summative evaluators commented on SLGs, and their comments primarily concerned the need for professional development that focused on helping teachers understand how to formulate their goals.

### *Summary of Student Learning Goals*

- MDE provided a *Student Learning Goals Handbook* and training to assist teachers with tasks related to developing their SLGs. Teachers' responses indicated that most of them found the Handbook and MDE training to provide adequate guidance.

- On average, classroom teachers spent 2.5 hours working on their SLGs while non-classroom teachers spent 1.7 hours. Summative evaluators reported that they spent an average of 50 minutes per teacher working on SLGs.

- Overall, 42% of teachers indicated that they found the setting of SLGs to be an easy task. To set their SLGs, teachers were required to select an assessment, establish mastery scores, and establish student start points. Of these tasks, the establishment of student start points was easier for the teachers than selecting an assessment or establishing mastery scores.

- Teachers were twice as likely to consult with a peer reviewer (69%) than with a summative evaluator (33%) in the development of their SLGs. Regardless of the individuals consulted, however, the majority of teachers reported that the consultations were beneficial.

- Summative evaluators tended to view SLGs more positively than classroom teachers did with respect to SLGs generating information to be used to assign a rating to a teacher's performance or to identify areas where a teacher might benefit from professional development. Approximately three-fourths of summative evaluators gave positive ratings to the use of SLGs for these two purposes compared to approximately 56% of classroom teachers. The opinions of non-classroom teachers were somewhat more positive, with over two thirds giving a rating of *good* or better.

# *Student Engagement*

The Model stipulates that results of a reliable and valid student survey will make up 15% of the student engagement evidence and an additional 5% will come from other measures of engagement (e.g., observations, videos) for a total of 20% of a teacher's performance rating.  If a survey is not available, as would likely be the case for teachers of lower primary grades, then the entire 20% will be made up of other measures of engagement.

## *State Model Student Survey Instrument*

To assess student engagement, MDE worked with a group of Minnesota educators and survey design experts to develop student perception surveys for use by the Minnesota districts participating in the Pilot. MDE also contracted with YouthTruth, a national nonprofit organization, to provide a platform for administering the student survey, to develop an online system for reporting student survey results, and to provide professional development on survey interpretation.

Two different versions of the student survey were created, one for students in grades 3 through 6 and the other for students in grades 6 through 12. The surveys were titled "Student Perception Survey of Student Engagement" and contained items designed to obtain students' perceptions of their classroom experiences regarding academic, behavioral, cognitive, and affective domains surveys (e.g., *Our class stays busy and does not waste time,* and *This teacher asks questions to be sure we understand*). In this report, the student survey is referred to as the State Model Student Survey.

## *State Model Student Survey Sampling Plan and Administration*

*Survey sampling plan*. The sampling plan developed by MDE, in consultation with YouthTruth, required districts first to identify the classes in which students would be taking the survey. Several selection criteria needed to be met. For example, the selection criteria for elementary schools stipulated that teachers' classes should be grade 3 or higher, class periods should have at least five students, and students should not be asked to respond to more than two surveys. The plan for middle schools and high schools stipulated that three class periods should be selected for each teacher whose students would be surveyed, students should not be surveyed in more than five classes, and class periods should have at least five students. The middle and high school sampling plan also stated that the structure of the survey administration would allow students to complete all surveys in one sitting. After selections were made, YouthTruth worked with district staff to export class roster data (e.g., class name, teacher ID, student ID, student demographics) for the chosen classes.

*Survey administration*. State Model Student Surveys were completed online, and each school was assigned to a "round" with specific launch and closing dates. YouthTruth provided log-in codes in advance of the survey launch for all students enrolled in the selected classes.

*Summative evaluator satisfaction with the student survey administration process*. The summative evaluators were asked to indicate how satisfied they were with the process of

administering the State Model Student Survey to the students in their buildings. Responses ranged from *very dissatisfied* to *very satisfied* (see Figure 21). Although the responses showed a great deal of variability, the most frequent response was *neither dissatisfied nor satisfied*, and respondents were more likely to select *very satisfied* than *very dissatisfied*. It appears, therefore, that the administration process was acceptable for most of the districts.
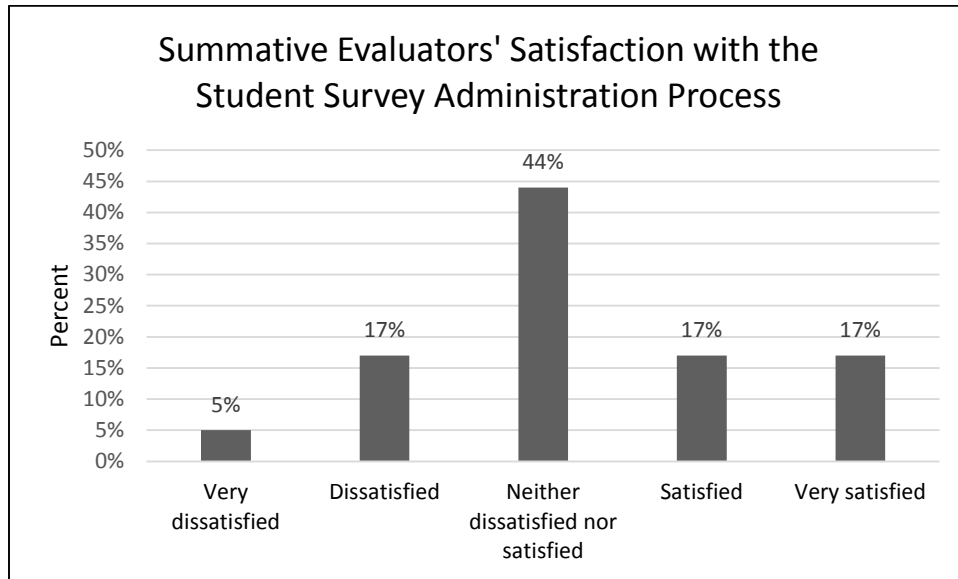


Figure 21.  Distribution of satisfaction ratings given by summative evaluators to the process of administering the State Model Student Survey.

Open-ended comments help explain the wide range of satisfaction levels regarding the survey administration process. Summative evaluators who were pleased with the process indicated that administration of the survey went smoothly. For example:

- *The process for administering the survey was not intrusive on class time and the students were able to complete the surveys on their own.*

On the other hand, summative evaluators expressing low levels of satisfaction mentioned that confusion resulted when students completed surveys for two or more teachers at one sitting. For example:

- *When doing more than one teacher survey, questions bounced back and forth between teachers, confusing students.*

Other dissatisfied respondents expressed concern that students did not seem to take the survey seriously. For example:

- *Even though students were prepped for the survey, it was common knowledge that students did not take the survey seriously to the extent that some took less than 2 minutes to complete the entire survey. Many created patterns with their answers.*

### *Training and Support*

The training and support provided by YouthTruth in collaboration with MDE and was made up of several components.

- A kickoff webinar for superintendents with an overview of the survey process and what to expect
- A survey implementation guide for participating schools, as well as phone and email support throughout survey implementation
- A live and recorded webinar for evaluators in the spring on how to interpret student survey data
- Follow-up resources for interpreting and discussing student feedback
- A video tutorial for teachers introducing the student feedback reports

*Summative evaluator satisfaction with training and support*. The summative evaluators were asked to indicate how satisfied they were with the training and support they received from YouthTruth. The response distributions are displayed in Figure 22. The summative evaluators were fairly equally divided between the response options of *satisfied* and *neither dissatisfied nor satisfied*, indicating that the summative evaluators perceived the YouthTruth training and support to be adequate.
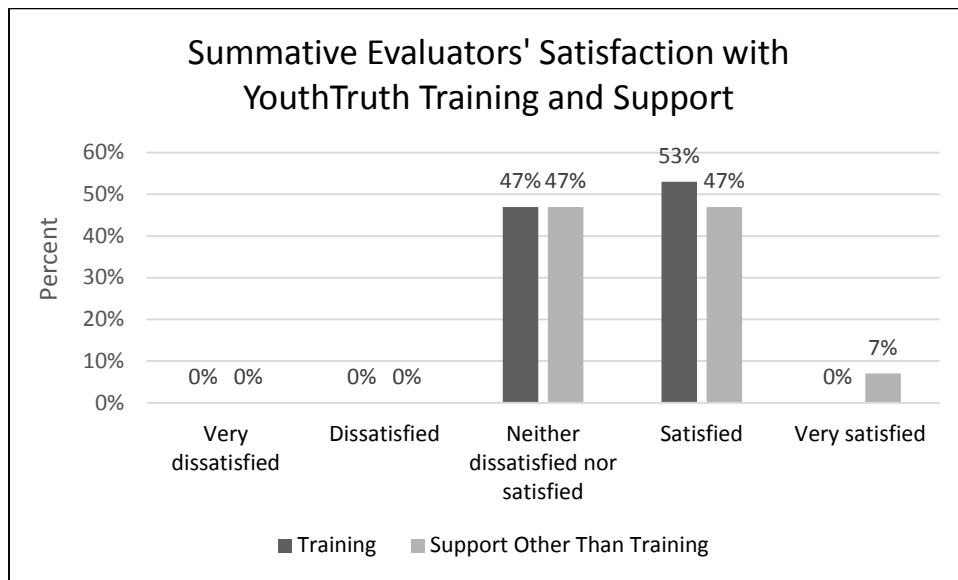


Figure 22.  Distributions of satisfaction ratings given by summative evaluators to YouthTruth training and support.

*Methods Used to Assess Student Engagement*

When appropriate, 15% of student engagement evidence is to be based on a student survey and 5% on other evidence. To find out the types of student engagement evidence that were utilized, teachers were asked to identify their student engagement assessment methods. For methods other than the State Model Student Survey administered by YouthTruth, a list of possibilities was presented (e.g., observation, video) and teachers were instructed to check all the methods they had used. An "other" response option was also provided where teachers were asked to write in any other methods they had utilized that were not on the list.

The methods used by classroom teachers to assess student engagement are presented in Table 10. Overall, 74% of classroom teachers said they assessed student engagement via the State Model Student Survey either by itself (6%) or in combination with some other method (68%), and 26% said they assessed student engagement only by a method other than the State Model Student Survey. The most frequently named other methods were observation (89%), student presentations, performances, and/or projects (48%), and surveys (27%).

Table 10.  Methods used by classroom teachers to assess student engagement

| **Methods used to assess student engagement (*n = 206*)** | **%** |
|---|---|
| State Model Student Survey in combination with other methods | 68% |
| Other methods only | 26% |
| State Model Student Survey only | 6% |
| **Methods other than the State Model Student Survey (*n = 186*)** **(by themselves or in combination with the** **State Model Student Survey)** | **%** |
| Observation | 89% |
| Student presentations, performances, and/or projects | 48% |
| Survey (other than the State Model Student Survey) | 27% |
| Student portfolios | 21% |
| Video | 11% |
| Other methods (e.g., discussions with students, exit slips) | 6% |
| Teacher portfolio | 5% |

*Time Spent by Teachers on the Assessment of Student Engagement*

Responses of classroom teachers regarding the total amount of time they spent working on the assessment of student engagement ranged from 0 to 20 hours with a mean of 2.90 hours ($SD = 4.05$, $n = 180$).  The most frequent response given by classroom teachers was 1 hour.

### *Time Spent by Summative Evaluators on the State Model Student Survey*

The summative evaluators were asked to indicate how many of the teachers for whom they served as summative evaluator had their students take the State Model Student Survey. Responses ranged from 0 to 55 teachers with a mean of 17.89 (*SD* = 15.76, *n = 18*). Including only the summative evaluators whose teachers had their students take the State Model Student Survey, the total time summative evaluators said they spent on activities related to the student survey ranged from 0 to 6 hours, with a mean of 2.37 hours (*SD* = 1.57, *n = 17*). The most frequent response was 2 hours, which was given by four summative evaluators.

### *Classroom Teachers' Opinions of Using Student Engagement Evidence for Teacher Evaluation*

On the teacher implementation survey, the respondents were asked to provide two ratings regarding student engagement evidence *in general* rather than rating a specific method such as observation or a student survey. Teachers were first asked to rate engagement assessments with respect to providing information that enables a summative evaluator to assign a valid rating to a teacher's performance. Next, they were asked to rate engagement assessments with respect to generating information that enables a summative evaluator to identify areas where a teacher could benefit from professional development. The two distributions, displayed in Figure 23, are remarkably similar, indicating that the teachers felt the same way about student engagement assessments, whether used for assigning a performance rating or for identifying areas for professional development. For both uses, 57% of the teachers chose a rating of *poor* or *fair*, and the most frequently chosen rating for both uses was *fair*. Thus, it appears that, overall, teachers do not enthusiastically support the use of student engagement evidence for their performance evaluations.

## Classroom Teachers' Opinions of Using Student Engagement Assessments for Teacher Evaluation

(Bar chart showing Percent on the y-axis from 0% to 35%, with categories Poor, Fair, Good, Very Good, Excellent on the x-axis)

| Category | To Assign Performance Rating | To Identify Areas for Professional Development |
|---|---|---|
| Poor | 27% | 25% |
| Fair | 30% | 32% |
| Good | 27% | 26% |
| Very Good | 12% | 13% |
| Excellent | 3% | 4% |

■ To Assign Performance Rating    ■ To Identify Areas for Professional Development

Figure 23.  Distributions of ratings given by classroom teachers to student engagement assessments with respect to their generating information that enables a summative evaluator to assign a valid rating to a teacher's performance and to identify areas where a teacher might benefit from professional development.

To investigate whether or not the teachers' ratings were related to the grade levels that they taught, teachers were placed in one of three mutually exclusive grade-level bands:  Grades K-2, grades K-5, and grades 6-12. More specifically, teachers were placed in the K-2 band if they taught one or more grades in the range kindergarten through grade 2 but did not teach any grades higher than grade 2. Teachers in the K-5 band taught one or more grades in the 3-5 range or taught one or more grades in the 3-5 range as well as one or more grades in the K-2 range. Teachers in the grade 6-12 band taught one or more grades in the range grade 6 through grade 12 but taught no grades lower than grade 6. All mean differences were analyzed via a Bonferroni planned comparison procedure with the type 1 error probability set at .05 for each set of pairwise comparisons. None of the pairwise comparisons resulted in a statistically significant difference. Therefore, the analysis results indicate that teachers hold the same opinion of using student engagement assessments for teacher evaluation, regardless of the grade levels that they teach.

### *Summary of Student Engagement*

- A State Model Student Survey was developed to assist districts with the assessment of student engagement. The summative evaluators indicated that they were satisfied with the process of administering the survey and with the training and support provided to the districts.

- The two most common methods utilized by classroom teachers to assess student engagement were student surveys—the State Model Student Survey or a different survey (100%) and observations (89%).

- The summative evaluators reported that they spent an average of 2.4 hours on activities related to the State Model Student Survey. The average time spent on assessment of student engagement by classroom teachers and non-classroom teachers was 2.9 hours and 1.6 hours, respectively.

- Teachers were asked their opinions about using student engagement evidence in their performance evaluations and their responses indicated that they were not enthusiastically supportive. More specifically, 57% of teachers gave a rating of *poor* or *fair* to the use of student engagement assessments for assigning a valid rating to a teacher's performance and for identifying areas where a teacher could benefit from professional development.

## *Overall Perceptions of the Model*

### *Benefits of the Model*

The teacher and summative evaluator surveys presented six possible benefits associated with the use of the Model and asked the respondents to check the ones on the list that, based on their experience, they perceived to be benefits. The results are summarized in Table 11.

Table 11.  Perceived benefits associated with use of the Model

| Based on your experience, what would you say are the benefits associated with the use of the Model? (Check all that apply) | % selecting the response option | | |
|---|---|---|---|
| | Classroom teacher *(n = 272)* | Non-classroom teacher *(n = 11)* | Summative evaluator *(n = 20)* |
| Facilitates teacher-to-teacher conversations regarding instructional practice | 68% | 100% | 75% |
| Facilitates evaluator-to-teacher conversations regarding instruction practice | 50% | 55% | 70% |
| Provides information on how teachers can improve their instructional practice | 40% | 64% | 70% |
| Enables teachers to see improvement over time | 39% | 45% | 45% |
| Incorporates multiple sources of information | 35% | 55% | 70% |
| Informs district decisions regarding professional development opportunities | 23% | 36% | 55% |

For all three survey participant groups, the most frequently selected benefit associated with use of the Model was that it facilitates teacher-to-teacher conversations regarding instructional practice. Another benefit frequently selected by all three groups was that use of the Model facilitates evaluator-to-teacher conversations about instructional practice. However, for all options included in the list, the percent of classroom teachers selecting the option as a benefit was less than the corresponding percent of non-classroom teachers and summative evaluators who selected the option. These results suggest that classroom teachers were less likely than non-classroom teachers and summative evaluators to see benefits associated with use of the Model.  In general, summative evaluators perceived the Model to have relatively more benefits than did the teachers or the non-classroom teachers.

### *General Overall Ratings of the Model*

The summative evaluator and the teacher questionnaires included two items that asked the respondents, based on their experiences, to rate the Model overall using a scale of *poor*, *fair*, *good*, *very good*, and *excellent*. One of the items asked the respondents to provide a general overall rating of the Model, and the other item asked them specifically to rate the Model's ability to provide a valid teacher performance assessment. The teacher and summative evaluator versions

of the questionnaire presented slightly different wordings of the latter item. The teacher version asked the respondents to rate the Model's ability to provide a valid rating of *your* performance, and the summative evaluator version asked the respondent to rate the Model's ability to provide a valid rating of *a teacher's* performance. For analysis purposes, the response options were assigned numerical values from 1 to 5 with *poor* = 1 and *excellent* = 5. Analyses were first carried out to compare the ratings given by classroom teachers, non-classroom teachers, and summative evaluators. Additional analyses were then conducted on only the classroom teachers' ratings.

*Classroom teachers, non-classroom teachers, and summative evaluators*. Distributions of the general overall ratings provided by classroom teachers ($n = 258$), non-classroom teachers ($n = 12$), and summative evaluators ($n = 21$) are displayed in Figure 24. The frequency bars of the summative evaluators' ratings are associated with noticeably more positive ratings than those of the classroom teachers and the non-classroom teachers. In addition, the results of analyses carried out on the group means indicate that the summative evaluators' mean rating ($M = 3.00$, $SD = 0.78$) was significantly higher than that of the classroom teachers ($M = 2.29$, $SD = 0.85$). Neither the classroom teachers' mean nor the summative evaluators' mean differed significantly from the mean rating of the non-classroom teachers ($M = 2.42$, $SD = 1.09$).
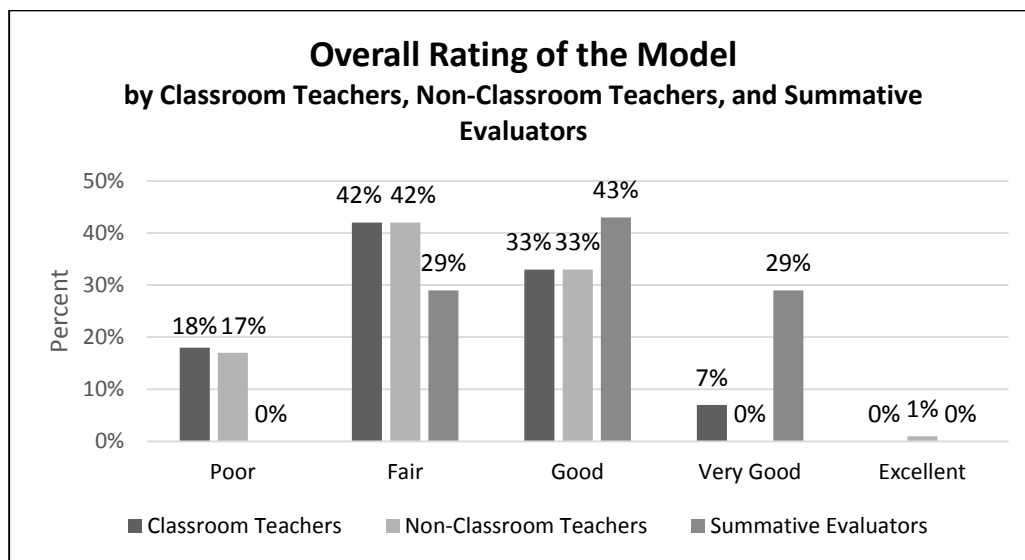


Figure 24.  Distributions of overall ratings of the Model given by teachers, non-classroom teachers, and summative evaluators

*Grade-level bands*. To investigate whether or not the teachers' ratings were related to the grade levels that they taught, the same procedure described earlier in the student engagement section was utilized to analyze grade-level band differences in overall ratings. The response distributions of teachers who taught students grades K-2 ($n = 34$), grades K-5 ($n = 46$), and grades 6-12 ($n = 133$) are displayed in Figure 25. The distributions of the three groups are quite similar with the highest frequencies associated with ratings of *fair* and *good*. None of the pairwise comparisons carried out on the mean ratings of the grade K-2, grade K-5, and grade 6-12 groups resulted in a statistically significant difference. We can, therefore, conclude that teachers' hold the same overall perceptions of the Model regardless of the grade levels that they teach.

**Overall Rating of the Model**
by Classroom Teacher Grade-Level Band
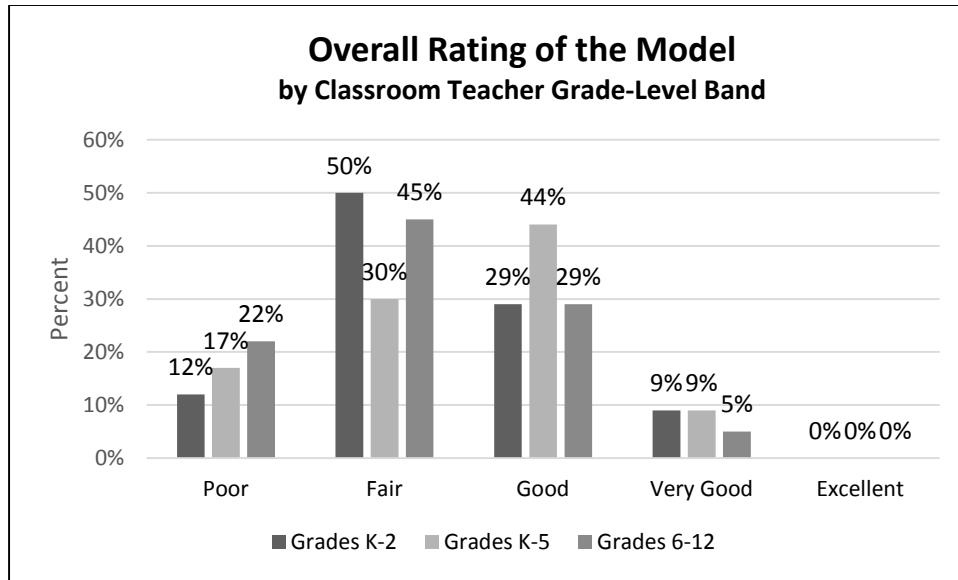
Figure 25. Distributions of overall ratings of the Model given by teachers in grade-level bands K-2, K-5, and 6-12.

### *Ratings of the Model for Generating Information That Enables the Summative Evaluator to Assign a Valid Performance Rating*

*Classroom teachers, non-classroom teachers, and summative evaluators*. Distributions of the ratings provided by classroom teachers ($n = 254$), non-classroom teachers ($n = 12$), and summative evaluators ($n = 19$) of the Model's ability to provide a valid assessment of a teacher's performance are displayed in Figure 26. Similar to the distributions of the general overall ratings, the frequency bars of the summative evaluators' distribution are associated with noticeably more positive ratings than those of the teachers. The results of the pairwise comparisons carried out on the group means indicated that the summative evaluators' mean rating ($M = 2.84$, $SD = 0.69$) was significantly higher than that of the classroom teachers ($M = 2.13$, $SD = 0.86$) and of the non-classroom teachers ($M = 2.08$, $SD = 0.79$). The mean ratings of classroom teachers and non-classroom teachers were not significantly different.
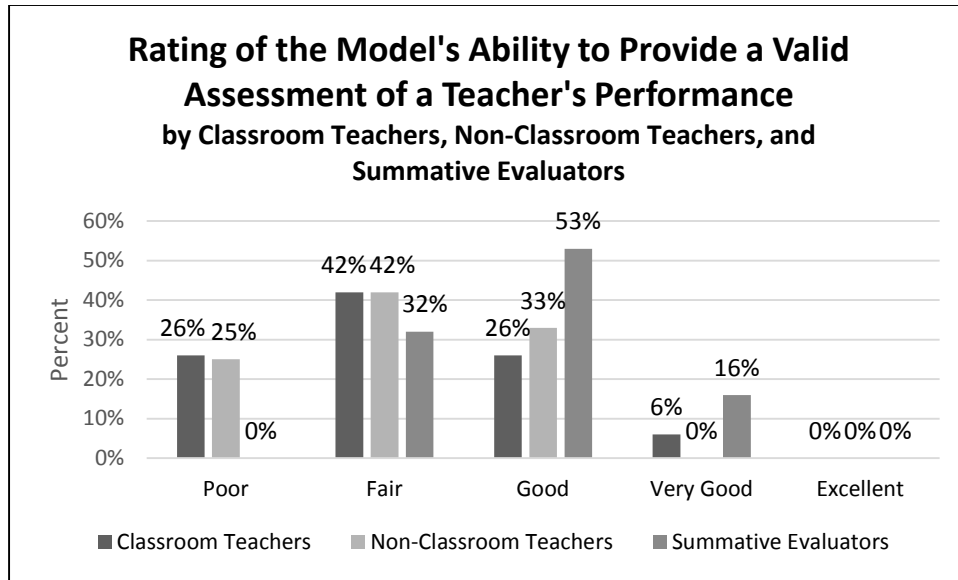
Figure 26. Distributions of ratings of the Model's ability to provide a valid assessment of a teacher's performance given by teachers, non-classroom teachers, and summative evaluators

*Grade-level bands*. The classroom teacher rating distributions for grades K-2 ($n = 35$), grades K-5 ($n = 44$), and grades 6-12 ($n = 132$) are displayed in Figure 27. In each grade-level band, the rating with the highest frequency was *fair*. Compared to the other grade bands, teachers of grades 6-12 appeared to be more likely to select a rating of *poor* and less likely to select a rating of *good*. Therefore, it was not unexpected that the mean rating of teachers in the grade 6-12 band ($M = 1.93$, $SD = 0.80$) was significantly lower than the mean rating of teachers in both the grade K-2 band ($M = 2.23$, $SD = 0.88$) and the grade K-5 band ($M = 2.32$, $SD = 0.86$). The mean ratings of the grade K-2 band and the grade K-5 band were not significantly different.



Figure 27. Distributions of ratings of the Model's ability to provide a valid assessment of a teacher's performance given by teachers in grade-level bands K-2, K-5, and 6-12.

*Change in Opinion*

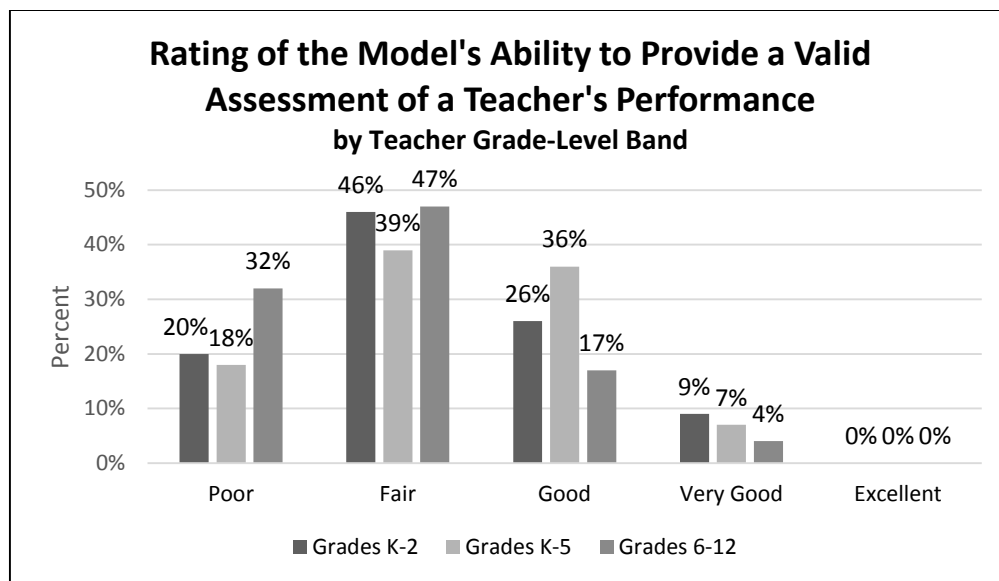Both the teacher survey and the summative evaluator survey included an item that asked, "*Now that you have acquired more experience working with the Model, has your opinion about the Model changed*?" If they responded "yes," they were asked to provide an explanation. The percentages of classroom teachers, non-classroom teachers, and summative evaluators responding "yes" were 24%, 17%, and 29%, respectively. Their explanations are summarized in the following paragraphs.

*Classroom teachers*. Fifteen of the explanations provided by teachers mentioned the time required to implement the Model as a drawback. The two examples presented below are representative of teacher comments that mentioned time.

- *It was unacceptable to expect busy teachers to spend so much time on this Model.*
- *Time was the biggest issue.*

Also mentioned fairly frequently ($n = 9$) by teachers was that they found the Model to be confusing or cumbersome. For example:

- *It was a good starting point, but it is very cumbersome and needs to be clearer in direction. All components must be given at the beginning of the training so teachers can see the entire picture. Getting one component at a time is difficult to plan ahead for the entire process.*
- *I found the various parts of it confusing and difficult to keep separate.*

On the other hand, some teachers ($n = 8$) did comment that their experience with the Model helped them to understand the evaluation process better. For example:

- *With more understanding came the ability to use it as a tool to improve.*
- *I have a greater understanding of the process.*

In addition, several teachers ($n = 12$) stated that they obtained helpful information by participating in specific components of the Model. The comments shown below describe peer observation, SLGs, and the IGDP as being valuable.

- *I was nervous at first, but the peer feedback helped my performance and the peer observation allowed me to adjust my performance as well.*
- *I thought it was like most MDE hoops; useless, confusing, and a big waste of time. But I actually found it beneficial to make a student learning goal for myself, as it helped lead discussion in our PLC meetings*
- *I found value in the IGDP and SLG.*

Six comments by teachers specifically mentioned the student survey, and these comments were equally divided between presenting positive and negative perceptions. An example of a positive and negative perception, respectively, are shown below.

- *The only good part of the teacher development is the student survey. This gives a good indication of what the teacher is doing in the classroom. That is the only part that I felt was valuable.*
- *Students really have a hard time being able to provide consistent feedback, students don't always take these things seriously.*

***Non-classroom teachers***. Only two respondents who were non-classroom teachers provided explanations as to why their opinion of the Model had changed and both explanations presented a positive view of the Model. One of these explanations is shown below.

- *This is a thorough tool. Initially, it was very intimidating, but with use it became more manageable.*

***Summative evaluators***. Four of the summative evaluators provided an explanation as to why their opinion of the Model had changed. In two of the explanations, the summative evaluators stated that their districts had decided to use their own teacher evaluation systems and that these systems incorporated some aspects of the MDE Model. In another explanation, a summative evaluator stated the student survey component worked very well. The only explanation provided by summative evaluators that was critical of the Model expressed frustration with the percentages applied to the components of the Model, but no details were given as to why the percentages were not satisfactory.

## *Recommendations for Changing the Model*

The survey presented five possible ways in which the Model might be changed to make it more effective, and respondents were asked to check all with which they agreed (see Table 12). When these options were ordered from most to least often selected, the ordering in each survey participant group was exactly the same. Namely, the top three recommendations selected by classroom teachers, non-classroom teachers, and summative evaluators were streamline, simplify, and clarify. Only about one fourth of the respondents felt that the component weights should be changed, and even fewer thought that the number of rating categories should be changed. It should be noted that the summative evaluators were especially in strong agreement that the Model was too time consuming and needed to be streamlined.

Table 12.  Recommendations for changing the Model so that it would be more effective

| How might the Model be changed so that it would be more effective?  (Check all that apply.) | % selecting the response option | | |
| --- | --- | --- | --- |
| | **Classroom teacher** *(n = 265)* | **Non-classroom teacher** *(n = 12)* | **Summative evaluator** *(n = 19)* |
| Streamline; it's too time consuming. | 60% | 67% | 95% |
| Simplify; it's too complex. | 56% | 50% | 63% |
| Clarify; it's difficult to follow the instructions or see how the parts fit together. | 49% | 42% | 42% |
| Change the component weights; current weights are .45 Teacher Practice, .20 Student Engagement, .30 Student Learning and Achievement | 25% | 25% | 26% |
| Change the number of rating categories; current categories are Exemplary, Effective, Development Needed, Unsatisfactory | 11% | 0% | 11% |

## *Open-Ended Comments*

After answering questions about specific aspects of the Model, the survey participants were invited to share any comments or suggestions they had regarding the Model. Open-ended comments were provided by 37 classroom teachers, 3 non-classroom teachers, and 3 summative evaluators. The content of their comments was analyzed to identify frequently occurring topics. These topics are summarized below.

### *Classroom Teachers' Comments:*

*Time needed to implement the Model*. By far, the topic mentioned most frequently by classroom teachers was the amount of time needed to implement the Model (*n* = 15). The teachers stated that the evaluation process was overly time consuming for their principals as well as for themselves. Two example comments are presented below.

- *How is a principal supposed to have points of contact with all of their staff and still manage discipline and other student needs at the school at the same time? It seemed from the beginning that the amount of time in the school year was not enough to meet the expectations that were laid out by MDE. I don't feel that it would be a sustainable way of measuring teacher evaluation.*
- *It was very time consuming. My time would have had more impact if I spent that time designing lessons and activities for my students.*

*Cumbersome/confusing*. Several teachers (*n* = 8) stated that they found the Model to be cumbersome or confusing. For example:

- *The intentions are good, but too many steps and hoops to jump through.*
- *In order for this to be done well, there needs to be experts on the Model at the schools. It seemed that no one really knew what was going on so we all just filled out the next form and moved on.*

*Student survey*. Seven teachers who participated in the student engagement component of the Model mentioned problems with the student survey, especially with respect to the questions on the survey instruments and the validity of the students' responses.

- *The questions were somewhat confusing in the YouthTruth survey. Phrasing a question for 3rd graders and seniors the same way gets you very different reactions.*
- *You could easily predict student responses based on the type of students you teach. I was rated higher across the board by my high achieving students. . . We should not continue using this survey to test student engagement.*

*Complete dissatisfaction*. A small number of teachers (*n* = 3) simply expressed complete dissatisfaction with the Model. For example:

- *Scrap it and start over. Seriously.*

*Positive outcomes*. Only a handful of teachers (*n* = 3) mentioned positive outcomes associated with implementation of the Model. The positive outcomes mentioned enhanced teaching and student success due to discussions with peers, and students being able to see their growth over time when teachers shared results from pre- and post-assessments with them.

*Non-Classroom Teachers' Comments:*

The comments made by the three non-classroom teachers concerned two topics. These were that too much time was required to implement the Model, and the Model did not fit non-classroom teachers very well.

### *Summative Evaluators' Comments:*

Only one topic was mentioned in the comments made by three summative evaluators and that was that the process was very time consuming. One summative evaluator pointed out that the time requirements were especially demanding for small districts where a building has only one administrator.

### *Summary of Overall Perceptions of the Model Findings*

- Summative evaluators' ratings of the overall Model were significantly more positive than those given by either classroom teachers or non-classroom teachers. The teachers' lower ratings appeared to be related mostly to the time required to implement the Model, aspects of the Model that teachers found to be confusing or cumbersome, and doubts about the validity of feedback received from the student survey.

- The benefits of the Model most frequently identified by both teachers and summative evaluators was that the Model facilitated teacher-to-teacher conversations and summative evaluator-to-teacher conversations about instructional practice.

- Teachers and summative evaluators were in strong agreement regarding the changes that should be made to the Model to make it more effective. Their top three recommendations, in order of highest to lowest, were: a) streamline the Model; it's too time-consuming, b) simplify the Model; it's too complex, and c) clarify the Model; it's difficulty to follow the instructions or see how the parts fit together.

# ♦ TEACHER BELIEFS ABOUT STUDENT ENGAGEMENT MEASURES SURVEY

The primary purpose of the teacher beliefs about student engagement measures survey was to investigate teachers' attitudes and beliefs regarding the use of student survey data in teacher performance ratings. Two main questions were addressed:

1. How did teachers feel about the use of student surveys before the Pilot was implemented in their districts?
2. Did their opinions change by the end of the process when the student survey results had been received and reviewed with the summative evaluators?

Analyses were carried out to answer additional questions regarding whether or not teachers' opinions differed by grade levels taught, subjects taught, or years of teaching experience. Teachers' and principals' opinions regarding the validity and usefulness of student surveys were also compared.

## Methodology

### *Instruments*

#### *Teacher Questionnaires*

Two questionnaires were developed to elicit teachers' opinions regarding the use of student surveys to evaluate teacher performance. The pre-implementation survey was administered at the beginning of the school year before activities related to the student survey were initiated, and the post-implementation survey was administered in late spring near the end of the school year after teachers were expected to have received a summary report of their students' survey results. In this report, these two teacher surveys are referred to as the fall survey and the spring survey, respectively.

On both the fall and spring surveys, teachers were asked to indicate their level of agreement with each of 12 statements regarding student surveys (e.g., *I am in favor of using feedback from student surveys to evaluate my teaching* and *I am concerned that students will give low ratings in a subject that is difficult for them)*. A seven-point response scale was utilized (*strongly disagree, disagree, somewhat disagree, neither disagree nor agree, somewhat agree, agree, strongly agree*). Teachers were also asked to provide background information regarding grade level(s) taught, subject area(s) taught, and number of years of teaching experience. On the spring survey but not the fall survey, teachers were asked how they would rate student surveys with respect to generating information for two intended outcomes of the Model: a) assigning a valid rating to a teacher's performance, and b) identifying areas where a teacher could benefit from professional development. Copies of the fall and spring questionnaires are provided in Appendix F and Appendix G, respectively.

*Summative Evaluator Questionnaire*

Two items on the summative evaluator questionnaire are included in analyses reported in this section. With the exception of minor differences in wording, these were the same two items presented on the spring teacher beliefs survey that asked respondents to rate State Model Student Survey with respect to generating information for assigning a valid rating to a teacher's performance and for identifying areas where a teacher could benefit from professional development.

*Procedure*

The fall teacher beliefs survey was administered as a paper/pencil instrument. The fall surveys were either hand delivered to the districts or sent by mail. Completed surveys were placed in envelopes, sealed, and returned. Spring surveys for both teachers and principals were administered online.

*Participants*

A total of 791 K-12 classroom teachers responded to the fall teacher beliefs survey, resulting in an overall response rate of approximately 86%. A total of 466 K-12 classroom teachers responded to the spring survey, resulting in an overall response rate of 51%. Forty-four of these cases were deleted because the respondents only provided answers to items requesting demographic information (e.g., district, subjects taught) and omitted all other items. Because analyses were carried out on matched pairs of individuals, the sample size reduced to 219 teachers who had complete data for both the fall and the spring questionnaires.

# Findings of the Teacher Beliefs Survey

## *Teachers' Opinions in the Fall*

Tables included in Appendix H display the percent of teachers selecting each response option. To address the question of how teachers felt about student surveys before the Pilot was implemented, the agreement rate for each item was calculated based on responses given on the fall questionnaire. For reporting purposes, the questionnaire items are presented according to whether they expressed support for the use of student surveys or expressed a concern. Agreement rates for items in these two categories are presented below.

### *Support for Student Surveys*

Six of the 12 questionnaire items expressed support for the inclusion of student surveys in a teacher evaluation system. From 51% to 78% of teachers agreed with the statements by selecting one of the three response options that indicated agreement (i.e., *somewhat agree, agree, strongly agree*). These items and their agreement rates are shown below.

- *Student surveys will provide me with information that will enable me to improve my effectiveness as a teacher.* (78%)
- *Student surveys will provide me with feedback that will enable me to identify strengths and weaknesses in my teaching methods.* (77%)
- *Feedback from student surveys can be used to track changes in my teaching performance over time.* (60%)
- *Surveying my students is worth the time it takes to administer the survey.* (55%)
- *I am in favor of using feedback from student surveys to evaluate my teaching.* (55%)
- *Feedback from student surveys can help my school identify areas where teachers need professional development.* (51%)

It should be pointed out that the most frequently chosen response option for four of the six items shown above was *somewhat agree*. In addition, only slightly more than half of the teachers (55%) indicated that they were in favor of using feedback from student surveys to evaluate their teaching. These results suggest that, at the beginning of the school year before implementation of the Pilot had begun, the teachers were expressing tentative rather than strong support.

### *Concerns about Student Surveys*

Six of the 12 items expressed a concern regarding the use of student surveys in a teacher evaluation system. Four of the items expressed a concern in a negative manner (e.g., *I am concerned that students will give low ratings in a subject that is difficult for them*), and two of the items expressed the concern in a positive manner (e.g., *Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors*). Overall, from 30% to 69% of teachers indicated agreement (*somewhat agree, agree,* or *strongly agree*) with the six statements, and the most common response selected for four of the six items was *somewhat agree*. The six items and their corresponding agreement rates are displayed below.

- *I am concerned that students will give low ratings in a subject that is difficult for them.* (69%)
- *I am concerned that students will use the survey to get back at teachers.* (61%)
- *It will take a lot of time for me to review the results of my students' surveys in order to make them useful.* (44%)
- *Student surveys are not worth administering because I can easily influence students so they give positive ratings.* (38%)
- *Responses that my students give on the student survey can accurately predict how well they do on learning performance measures.* (35%)
- *Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors.* (30%)

Based on the levels of agreement associated with these items, it appears that teachers' support for the use of student surveys was mixed with reservations related to uncertainty about the students' ability to provide valid feedback and the time it would take to administer the student survey.

## *Change in Classroom Teachers' Opinions*

The fall and spring teacher belief questionnaires contained the same 12 items designed to solicit teachers' opinions about using student surveys in a teacher evaluation system. Ratings given at the beginning of the school year were compared to the ratings given by those same teachers near the end of the school year. Results are summarized in Tables 13 and 14 in terms of mean ratings.  The six items displayed in Table 13 are statements that express support for the use of student surveys for teacher evaluation and/or mention a benefit associated with student survey feedback. The six items displayed in Table 14 are statements that express concerns about the use of student surveys. When interpreting these results, it should be kept in mind that 90% of teachers of grades 3 and above whose students had taken the student survey indicated on the spring survey that they had seen their student survey feedback reports. However, only 28% said they had discussed the survey results with their summative evaluator.

Table 13. Fall to spring change in teachers' ratings of statements that express support for the use of student surveys for teacher evaluation

| Questionnaire item | $n$ | Fall to spring change | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | Mean rating[a] | % Selecting lower rating | % Selecting higher rating | $z$ | $p$ |
| I am in favor of using feedback from student surveys to evaluate my teaching | 218 | $M_{Fall} = 4.42$  $M_{Spring} = 3.92$ | 49% | 28% | 4.16 | .000 |
| Student surveys will provide me with feedback that will enable me to identify strengths and weaknesses in my teaching methods | 219 | $M_{Fall} = 5.02$  $M_{Spring} = 4.48$ | 46% | 22% | 4.42 | .000 |
| Student surveys will provide me with information that will enable me to improve my effectiveness as a teacher | 219 | $M_{Fall} = 5.09$  $M_{Spring} = 4.50$ | 46% | 16% | 5.66 | .000 |
| Feedback from student surveys can be used to track changes in my teaching performance over time | 219 | $M_{Fall} = 4.19$  $M_{Spring} = 3.79$ | 47% | 23% | 3.67 | .000 |
| Feedback from student surveys can help my school identify areas where teachers need professional development | 219 | $M_{Fall} = 4.53$  $M_{Spring} = 4.01$ | 49% | 25% | 4.47 | .000 |
| Surveying my students is worth the time it takes to administer the survey | 217 | $M_{Fall} = 4.43$  $M_{Spring} = 4.03$ | 47% | 26% | 3.29 | .001 |

[a]The 7-point rating scale was strongly disagree = 1, disagree = 2, somewhat disagree = 3, neither disagree nor agree = 4, somewhat agree = 5, agree = 6, strongly agree = 7.

## Statements Expressing Support for the Use of Student Surveys

All six items presenting support for the use of student surveys were associated with a statistically significant fall-spring difference where the mean agreement rating in the spring was lower than the mean agreement rating in the fall (see Table 13). Nearly half of the teachers selected a lower rating in the spring than they had in the fall. The content of these items indicates that, after teachers became more familiar with the student surveys, they were more tentative in their support of such surveys in a teacher evaluation system and more skeptical regarding the usefulness of student survey feedback. For example, the proportion indicating they were in favor of using feedback from student surveys to evaluate their teaching declined from 55% to 45%, and the proportion indicating they felt that student surveys would provide them with information that would enable them to improve their effectiveness declined from 78% to 67% (see Appendix H).

Table 14. Fall to spring change in teachers' ratings of statements that express concerns regarding the use of student surveys for teacher evaluation

| Questionnaire item | $n$ | Mean rating[a] | Fall-spring change | | $z$ | $p$ |
| --- | --- | --- | --- | --- | --- | --- |
| | | | % Selecting lower rating | % Selecting higher rating | | |
| I am concerned that students will give low ratings in a subject that is difficult for them | 217 | $M_{Fall} = 4.86$ $M_{Spring} = 5.11$ | 26% | 38% | 2.21 | .027 |
| I am concerned that students will use the survey to get back at teachers | 217 | $M_{Fall} = 4.63$ $M_{Spring} = 4.86$ | 24% | 37% | 2.51 | .012 |
| Student surveys are not worth administering because I can easily influence students so they give positive ratings | 219 | $M_{Fall} = 3.62$ $M_{Spring} = 3.86$ | 32% | 42% | 2.23 | .026 |
| Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors | 215 | $M_{Fall} = 3.88$ $M_{Spring} = 3.54$ | 44% | 27% | 3.32 | .001 |
| Responses that my students give on the student survey can accurately predict how well they do on learning performance measures | 217 | $M_{Fall} = 3.73$ $M_{Spring} = 3.38$ | 42% | 27% | 3.22 | .001 |
| It will take a lot of time for me to review the results of my students' surveys in order to make them useful | 217 | $M_{Fall} = 3.90$ $M_{Spring} = 3.47$ | 45% | 25% | 3.36 | .001 |

[a]The 7-point rating scale was strongly disagree = 1, disagree = 2, somewhat disagree = 3, neither disagree nor agree = 4, somewhat agree = 5, agree = 6, strongly agree = 7.

*Statements Expressing Concerns about the Use of Student Surveys*.

 All six items that expressed a concern about the use of student surveys in teacher evaluation were associated with a statistically significant fall to spring difference (see Table 14).  For the top three items in Table 14, the teachers expressed more agreement, and for the bottom three items, the teachers expressed less agreement. The content of the statements for which they expressed more agreement dealt with the validity of the survey with respect to the factors of course difficulty, students using the survey to getting back at teachers, and teachers being able to influence students to obtain positive ratings. The content of two of the statements for which they expressed less agreement dealt with students being able to discriminate between effective and ineffective teaching behaviors and the relationship between students' learning performance measures and their survey responses. The other item concerned time—the time needed to review the survey results in order to make them useful. The fall-spring change regarding time indicates that, after teachers had received their student survey results, they felt the review took less time than they initially thought it would.

### *Comparison of Classroom Teachers' and Summative Evaluators' Opinions*

Classroom teachers and summative evaluators were asked to provide two ratings specifically regarding the use of the State Model Student Survey for teacher evaluation. They were first asked to rate the State Model Student Survey with respect to its providing information that enables a summative evaluator to assign a valid rating to a teacher's performance (see Figure 28). They next were asked to rate the State Model Student Survey with respect to its helping a summative evaluator identify areas where a teacher could benefit from professional development (see Figure 29). In both graphs, the summative evaluators' ratings appear to be generally more positive than those of the classroom teachers.



Figure 28.  Distributions of ratings given by classroom teachers and summative evaluators to the State Model Student Survey with respect to its generating information that enables a summative evaluator to assign a valid rating to a teacher's performance.
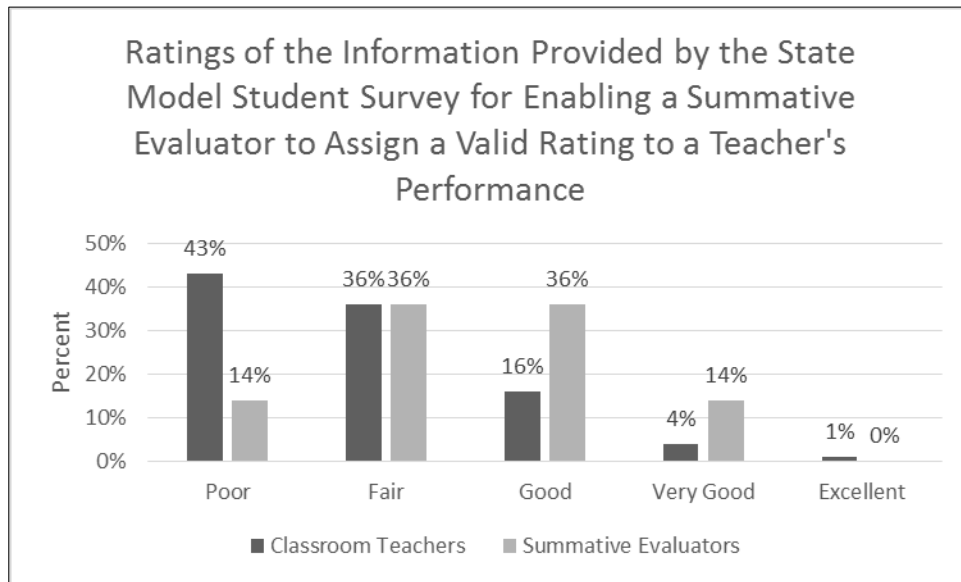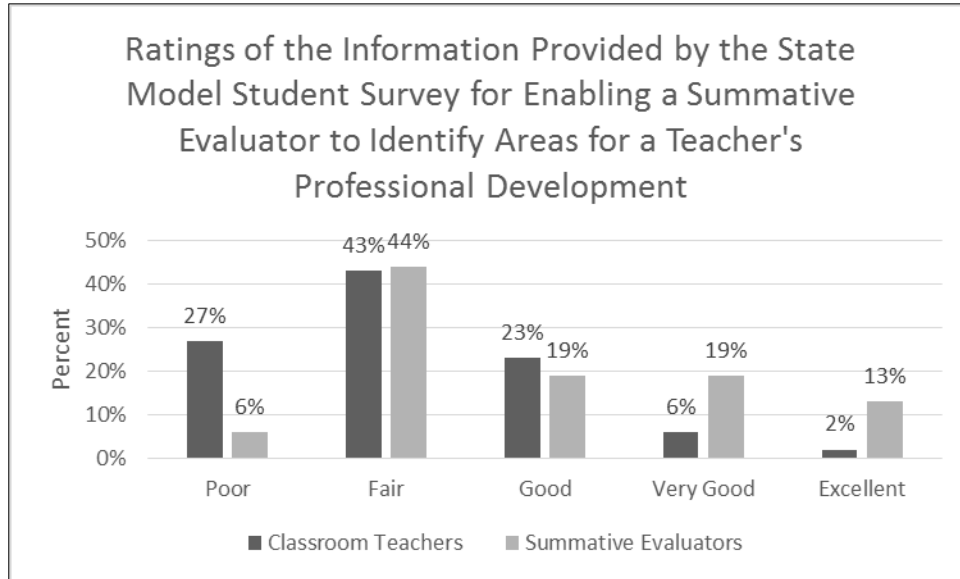
Figure 29.  Distributions of ratings given by classroom teachers and summative evaluators to the State Model Student Survey with respect to its generating information that enables a summative evaluator identify areas for a teacher's professional development.

The results of analyses carried out on the group means indicated that the summative evaluators' mean rating ($M = 2.50$, $SD = 0.94$) was significantly higher than that of the classroom teachers ($M = 1.84$, $SD = 0.90$) regarding use of State Model Student Surveys for assigning a teacher's performance rating. The summative evaluators' mean ($M = 3.33$, SD = 1.75) was also significantly higher than the classroom teachers' mean ($M = 2.14$, $SD = 0.94$) regarding use of State Model Student Surveys for identifying areas where a teacher could benefit from professional development. These results suggest that the summative evaluators see more value in the use of the State Model Student Survey data than classroom teachers with respect to assigning performance ratings as well as identifying areas where teachers might benefit from professional development.

***Summary of the Teacher Beliefs Survey Findings***

- At the beginning of the school year, before the Pilot was implemented in their districts, teachers indicated that they perceived student feedback to be potentially valuable for providing information that would help them improve their effectiveness.  More specifically, 55% of teachers said they were in favor of using student surveys to evaluate their teaching, and 78% said they believed that student survey feedback would enable them to improve their teaching effectiveness.

- On the survey administered at the beginning of the school year, teachers also indicated that they were uncertain about their students' ability to provide valid feedback. The most common concerns expressed by teachers on the fall survey were that students would use the survey to get back at teachers (61%) and that students would give low ratings in subjects that were difficult for them (69%).

- Near the end of the school year, after student survey feedback reports had been received, the teachers were significantly less supportive of including student surveys in their evaluations and significantly more concerned about the validity and usefulness of student survey feedback. Although their support had decreased, however, the teachers had not become strongly opposed to student surveys. On the spring survey, 45% of teachers indicated they were in favor of using student survey feedback to evaluate their teaching and 67% indicated they felt student surveys would provide them with information that would enable them to improve their effectiveness.

- When compared to teachers, the principals seemed to place more value on feedback from student surveys, both for the purpose of assigning a rating to a teacher's performance and for identifying areas where a teacher could benefit from professional development.

# ♦ TEACHER INTERVIEWS

## Methodology

### *Interview Protocol*

The interview protocol for teacher interviews was divided into the following sections: General, Teacher Practice, Student Engagement, Student Learning Goals, and Whole Model. There were 12 questions on the interview protocol, which is included in Appendix J.

### *Participants*

At the end of the 2013-2014 school year, CAREI evaluators interviewed 37 teachers across eight districts. Six of the districts were piloting the whole Model; one district was pilot testing student engagement and student learning goals; and one district was piloting the teacher practice component. The 37 teachers had also been interviewed by CAREI in the fall and had been selected by their superintendents and/or principals following criteria specified by CAREI. Of the teachers interviewed, 16 were elementary and 21 were secondary, with representation from both specialist teachers ($n = 16$) and core academic or class teachers ($n = 21$). Half of the specialist teachers ($n = 8$) were special education or teachers for English language learners (ELL), and the other half taught art, physical education, music, or family and consumer science. Just over one third of the teachers interviewed were probationary.

## Findings of the Teacher Interviews

### *Teacher Practice*

Seven of the eight interview districts participated in the teacher practice component, and 30 teachers from these districts were asked questions about this component. The questions concerned Individual Growth and Development Plans (IGDP), performance standards, points of contact with peers, points of contact with summative evaluators, teacher evaluations, and self-assessments. The teachers' responses are summarized below.

#### *Individual Growth and Development Plans (IGDP)*

An overwhelming majority (75%) of teachers found the IGDP to be very useful, with the remaining teachers expressing more neutral feelings about the process. None of those interviewed gave negative responses regarding the IGDP. The process of developing the IGDP seemed clear to teachers, partly because many of them had done something similar in the past. However, teachers described at least four different approaches to setting their professional goals: 1) some teachers were asked to align their individual professional goals with district or school goals, 2) some were grouped into new Professional Learning Communities (PLCs) according to their goal, 3) some decided on common goals within an existing or new PLC, and 4) some developed completely individual professional goals. This variety did not influence the quality of

the experience for the teachers interviewed, however, as satisfaction was evident across all approaches to the process.

The most common reasons expressed about why the IGDP worked for teachers was that it kept them focused.

- *As a teacher, so much comes at you. Let's do this, let's do that. If you have that goal and benchmark times to review it, it's a good thing.*
- *It was really nice to . . . sit down and think about what I wanted to do this year.*
- *It helped me focus in on how can I hold myself accountable? How can I use measurements and benchmarks and all that kind of stuff to really see if I accomplish my goal of being a better teacher or if it's just empty words?*
- *I think we need to get this so it's one of those things we see every single day, so it's out there and you kind of don't have to go back and look at what you wrote.*

Reasons for less enthusiastic responses included dissatisfaction with their professional goal, such as setting a narrow goal that was too easily achieved. Some teachers also found that their goals did not work well throughout the entire the school year.

- *I think the goals were a bit constricting as the year went on.*
- *Mine changed throughout the year. My original goals, I didn't keep either of them.*

### *Performance Standards for Teacher Practice Rubric*

Perceptions about the performance standards rubric were decidedly split between those who found it useful and those who did not use it at all. Only one of the seven districts was uniformly positive in talking about the rubric; all others were divided among themselves. Teachers who viewed the rubric positively spoke about its stability:

- *A lot of things in education change and change. This one seems to be stable and I think the categories are good. People are becoming familiar with them and understand them.*

Despite the mixed response, a number of teachers stated that they felt comfortable with the rubric from their experience with Danielson's *Framework for Teaching*. Additionally, the rubric was viewed as somewhat adaptable; one teacher described it as "armature" in a sculpture, something that might serve as a structure on which to build and grow.

- *I found it helpful to go through the performance standard sheets and actually look at the rubric and then go back and say: "Yup, those are the areas that I definitely need to grow in," and then to tie that into my planning. It really made me reflect on my teaching.*

On the other hand, several teachers believed that the rubric limits thinking and leads to "canned" ways of reflecting about their work.

- *Sometimes when you look at a rubric and you're making your Plan, you don't think. You just see what's on the rubric and you don't necessarily think outside the box. Sometimes too much guidance leads to canned answers.*

Still others only consulted it briefly, if at all.

- *I knew it was there and I guess I kind of glanced at it.*

### *Points of Contact with Peers*

Points of contact, according to the Model, may be formative or summative, depending on whether or not a peer reviewer or a summative evaluator is conducting them. The points of contact with peers mentioned by teachers were usually observations of their classroom practice or face-to-face meetings. While they found both the observations and the meetings to be helpful, they had quite a bit to say about how the points of contact were carried in their school and/or district, and offered ideas for improvements in the future.

Twenty-one out of 30 respondents (70%) commented specifically on the benefit of increased collaboration and stronger relationships among faculty resulting from the points of contact. Teachers in four of the seven districts completed most points of contact within their PLCs, rather than visiting each other's classrooms. Those who did classroom observations voiced the challenge of leaving their own students even if a substitute teacher was offered for a period of time, partly because planning for a substitute often results in a disruption to the instruction, and partly because (in two cases) teachers had only one day of warning before the substitute was made available. On the other hand, teachers also commented on how much they learned from each other when they were able to visit one another's classrooms:

- *I was flabbergasted with how much, and with what the other teachers are doing. It was quite an eye-opener!*
- *Anytime you can get the opportunity to build shared ideas with your fellow teachers— you're learning something. Too often, we all know, we're too busy in our own little worlds and we don't get out there to do that.*
- *We liked being able to go in and observe each other and get ideas from each other.*

Many points of contact with peers were carried out as informal or formal conversations such as discussions occurring during PLC meetings. The classroom observations typically also included professional discussions in the form of pre- and post-conferences. While the points of contact were generally viewed as useful, teachers found them terribly time-consuming when it came to documentation. The paperwork was thought to be unreasonable and repetitive by 12 out of 30 teachers.

- *Why do we have so many forms?*
- *It was a lot better just when we sat and talked about it. The value was in the discussion.*
- *Honestly, to sit down after the point of contact and get all the paperwork done was a lot. And you'd already talked about it verbally and then to go back and reiterate everything you just said and type it all—that was hard.*

Another concern for teachers was the fact that points of contact could be (and were) understood and enacted in multiple ways, even within a school. Several teachers spoke about hopes for clarity and consistency in how their school would implement the points of contact in the future. Much depended upon the peer with whom teachers were paired, and the perceived level of commitment of their partners. Not surprisingly, some teachers dove deeply into points of contact with peers and others went with a more abbreviated interaction. However it was carried out, they saw the ultimate goal as helping each other become a better teacher:

- *The spirit of this is the two of us sitting down together and me picking your brain or you picking my brain on how we can continue to become better teachers.*

While difficulties resulted from how the peer reviewers were grouped (within or across grade levels and disciplines), in several districts the difficulties with points of contact resulted from each reviewer having up to three "reviewees," causing them to feel overwhelmed. Multiple teachers in two districts felt it would be best to train all staff to be peer reviewers, which would distribute the workload more evenly.

- *I think it would be a whole lot easier if everybody got trained and everybody had maybe one person to coach/mentor. It would be easier because it would take less time, and easier to be in the role of reviewer if everyone experienced both sides.*

Another source of "pain" (a word used by two teachers in different districts to describe the points of contact) was that some teachers did not feel comfortable being honest with their peers about what they saw in the observations. In situations where peer reviewers were reluctant to give honest feedback, collegial relationships did not flourish despite increased contact between teachers. Even when honesty was not presented as a challenge, teachers at three districts talked about the need for clarity regarding the roles. Was it coach/mentor or evaluator/reviewer?

- *Going forward, they really need to differentiate between peer coaches versus peer evaluators, because that is going to be the part that causes heartache and pain for a lot of teachers. It's the part that's most uncomfortable at first. It took almost an entire school year to get most of the staff to the point of understanding that it's not people evaluating each other, it's people helping each other get better.*

### *Points of Contact with Summative Evaluators*

The points of contact completed by summative evaluators were commonly viewed as being part of a summative evaluation experience, rather than the more formative nature of points of contacts with peer reviewers. Teachers talked about varying levels of quality in the points of contact with summative evaluators. Key factors included the amount of time allotted to the pre- and post-observation conferences and the specificity of the feedback. Teachers appreciated when the points of contact were focused and specific, and several attributed this to the use of the IGDP and the rubric:

- *There were specific things* [the summative evaluator] *was looking for, and I knew what they were. The rubrics were right there. That was one of the best observations I've had—not meaning that I was rated high, but it was the most effective.*

Teachers in several districts thought summative evaluators could use more training on how to conduct classroom observations and follow-up conversations or specialists should be brought in to carry out the observations.

- *Summative evaluators need the training, too, and I don't always know if the principals are the best ones for that. I sometimes think they are so overwhelmed with staffing and IEPs and everything else that districts should look at actually hiring people to come in and do observations.*

Seven of the 30 teachers were worried about the time demands on summative evaluators. They were aware of the number of points of contact that summative evaluators were required to complete, and noted the difficulty this must have presented.

### *Evaluation Process*

In the interview group, 14 out of 30 teachers were given a full evaluation this year. Of these, only three stated unequivocally that they had received a numeric component rating for teacher practice in their final meeting with a summative evaluator. Several others were not sure if they had been given a numeric rating but thought it was possible. On the whole, teachers (13 out of 14) were fairly satisfied with the evaluation process. Several mentioned the fact that evaluations seemed more specific and less ambiguous due to the IGDP and the rubric.

- *The evaluation was more effective because it related directly to the different areas of teaching. I knew what I was being evaluated on and there weren't any surprises.*
- *For me as a veteran teacher, it* [teacher practice evaluation] *gave us something to critically look at in a positive way. It gave us things that we could talk about that maybe I'd want to improve on. I guess it gave us a tool for conversation.*

### *Self-Assessment*

Eighteen of 30 teachers completed the self-assessment. Twelve of 30 teachers did not complete or had no knowledge of the self-assessment form.

An overwhelming majority of teachers (94%) who completed the self-assessment thought the practice of "sitting down" to look at data at the end of the year was powerful, and they were happy to go through the exercise of reviewing all of the documentation they had generated. It took different amounts of time in different districts; whereas one district completed the self-assessment during a final staff meeting, teachers in another district described the process as extremely time-consuming. These teachers had multiple mentees, so they repeated the self-assessment process up to four times (including once for their own reflective practice). This group described completing the form, holding a meeting, and then documenting the meeting in their

software program, which was described as a "Mobius loop" because it seemed to fold back on itself unendingly. Teachers with multiple mentees, then, did not find the self-assessment experience valuable relative to the effort expended.

- *Was I able to put all the thought into it that I probably should have? Probably not. Actually, no, I didn't. I did not find it very beneficial because I was just restating everything that I had already written into the other part. Most of it was already commented on and then you had to restate your comments again. That really wasn't very valuable.*
- *Again, when you have to put it to paper, you can articulate what needs to be done, what needs to be learned from, what strengths, what weaknesses. That's always good.*

### *Student Learning Goals (SLG)*

Seven of the eight interview districts participated in the student learning and achievement component that contains the student learning goals (SLG) element. Thirty-three teachers were interviewed about SLGs and were asked questions regarding how they established and assessed class and targeted group goals, the utility of SLGs, and the challenges they encountered.

### *Class Goals*

Over 45% of the participating teachers used an internal measure to set their class learning goal, whereas just over one third used an external measure such as *MCA* or *BAS* test data. Most of the remaining teachers (15%) also used internal measures, but they sought "improvement" rather than specific quantified results. One teacher in the interview group did not set a student learning goal.

Special education teachers found the SLGs to be redundant, given the fact their students already have individual learning goals. Another wrinkle for special education teachers is that the Individual Education Plan (IEP) review process occurs throughout the year, depending on when it was created, so the SLG calendar for the Model might be quite out of synch with the IEP cycle.

Specialist teachers (art, physical education, family and consumer science, music, etc.) found the SLGs somewhat cumbersome, not because they did not want to set goals for their students, but because of complications due to subject matter, large numbers of students, and few meetings with all students in any given week.

- *How do you describe "Did they do well?" It's hard to gauge student success in a subject that isn't cut and dry. Was I happy with their progress? Yes. Can I show it in numbers? No.*

Roughly half of the teachers interviewed felt their students had met the class learning goal, although only 6 out of the 33 (18%) stated this unequivocally and had supporting data. Most of the teachers without supporting data who felt the class learning goal had been achieved spoke about being satisfied with the progress of their students; they used more general ways of assessing outcomes.

- *Everybody made growth and when I met with* [the summative evaluator] *that was the big thing. Like okay, two of my kids didn't quite make it where I said they were going to make it, but they still made growth.*

Of those who were not satisfied, quite a few wondered if they had "overreached" by setting the goal too high.

- *I did not meet my goal. I was a little disappointed on how low some of these kids scored. I just wonder if my goal was too high. That to me is the most challenging thing, creating a smart goal, I guess, where it's obtainable, and you're actually looking for growth.*

Conversely, one teacher who had set an "attainable" goal (and reached it) had questions from the summative evaluator about whether or not the goal was rigorous enough.

### *Targeted Group Goals*

Two thirds of the teachers interviewed set targeted group goals, while one third did not designate a targeted group. Of those who set a targeted group goal, 30% felt their students attained the goal, or almost attained it. Another 30% were not sure about the results for their targeted group for a variety of reasons, and 10% noted their targeted group had not reached the goal. The reasons for not setting a targeted group goal were mostly connected to the students that the teachers served. In the case of special education students, for example, all students might be considered "targeted." Other reasons for less vigorous participation in setting a targeted group goal may have been related to district implementation; teachers in three of the seven districts seemed uniformly vague when questioned about the targeted group goal. In fact, teachers from only one of the seven districts were specific in talking about the targeted group goal. The teachers in this district all found it very useful to have both class and targeted group goals, while the teachers in the remaining six districts were neutral or unenthused about drawing distinctions between groups of students for the SLGs.

- *My focus was on the kids having the most trouble.*
- *It was pretty clear who I need to meet with more, since it was right in front of me all the time.*

### *SLG Utility*

The teachers who found SLGs useful (40% of those interviewed) talked about the way the goals increased their focus, planning, and reflection.

- *The goal drove me because it had numbers associated with it.*
- *What are the skills we want kids to have, and how do we assess them? Then throughout the year we kept coming back to "How are we doing?" It was really our first year of getting into it at that level. We've always done it, but not focused and coordinated like we did this year.*

*SLG Challenges*

Teachers who commented on problems with the SLGs (45% of those interviewed) found them unwieldy for a number of reasons, not least that the SLGs were confusing:

- *I feel like that piece is still a shot in the dark for me.*
- *That part, of all the Model, that part was the least impactful for me because it didn't totally make sense for my situation.*
- *We need to make the student learning goals easier for the teacher to measure. To get data and change your practice based on that data—I'm not sure how to do that so I just did my best.*

Seven teachers (20%) talked about the fact they did not have an adequate number of students to warrant both class learning goals and targeted learning goals.

Four teachers (12%) felt the data didn't represent a complete picture of their work with students; one of these, while successful with the SLG data, noted there could be an element of "cherry-picking" to make the outcomes appear more positive. Some teachers pointed to tension between frank reporting of student progress and reporting data that will result in favorable evaluations.

- *It really seemed like I was going through the motions, just putting numbers to them* [student learning goals] *and it didn't mean a lot.*

### *Student Engagement*

Seven of the eight interview districts participated in the student engagement component, and 33 teachers were asked questions about this component. Interview questions concerned their perceptions of the State Model Student Surveys, including the perceived impact and problems that were encountered. They were also asked for recommendations to improve the student engagement component.

*General Perceptions*

Generally, teachers were open to the idea of student engagement surveys, but many pointed to problems with fairness. Of the teachers interviewed, 40% were positive about the survey experience, and a few more (45%) thought a similar survey should be used in the future.

- *The process was very, very rewarding for us here.*
- *I think it's good for students to take the same survey for all of their teachers, in addition to whatever they do in their classes.*
- *We're constantly evaluating ourselves as far as thinking about what's working, and what's not working? The survey helped me see if I need to adjust things because maybe I'm not communicating them right.*

Positive responses tended to touch upon the importance of hearing students' voices:

- *It's interesting to get the students' perspectives instead of another adult, or yourself.*
- *It was really positive. . . I think a lot of times we forget we should be catering to our students' learning, what type of learners they are and things like that. So if they're not engaged, they're not going to learn. And they really liked the chance to critique or grade their teachers.*
- *There's value in these kids knowing if they feel cared about, if they engage with the teacher, if they believe that their teacher is out for their best interest. I believe these are real feelings.*

However, quite a few teachers (36%) felt strongly that the survey should be modified or eliminated as part of the evaluation.

- *They were a nightmare.*
- *It was awful.*
- *It was hard not to take it personally.*

### *Impact of the Student Engagement Survey*

Almost half of the teachers interviewed cited self-awareness or planning for future practice as one outcome of the student engagement survey.

- *I really liked how you could break down to see differences between male and female responses.*
- *I changed like literally within a week after looking at the data. Why was I letting this happen? Sometimes you just have to take a step back and say: "I can do better."*

Teachers also commented on the fact they would want more data to see long-term trends.

- *We talked about how interesting it would be to see over time, not just one year's worth of results, but over time, trends that we would see.*

The remaining teachers (again, around half) did not find the surveys to have a discernible impact on their work.

### *Problems with the Student Engagement Survey*

Fourteen of the 33 teachers (42%) expressed concerns about the validity of the survey data, citing the age of their students as potentially affecting their ability to comprehend the survey items or value the survey's purpose.

Eleven teachers across the districts raised questions about who should take the survey. They were divided, however, in wondering what ages were likely to result in valid student engagement data:

1. Elementary students might not understand what was being asked.

- *I'm not really sold on surveying elementary kids. I just don't think they get it.*

2. Middle or high school students might try to sabotage a teacher's rating in a coordinated effort.

- *I know that students in one class got together to give low scores to a teacher they didn't like, but this teacher is an incredible teacher.*

3. Special education or ELL students might not understand the questions.

- *I had to read the questions out loud and modify the whole survey so my students would understand, and still I wasn't sure if they got it.*
- *I was in one room when kids were taking it and they were just kind of like "What does this mean?" and "What does that mean?" and I just said "We can't help you with it."*

The remaining three teachers raised questions about the timing of the survey. If given early in the semester, for example, teachers who had new classes would not know their students as well as teachers who had been with the same group for several months.

An additional area of concern was the way the State Model Student Survey data compared teachers against one another. This seemed unnecessary to many of the teachers interviewed, and at least one teacher at every interview district expressed concern for their colleagues in the event of low student survey results. They seemed to be as worried about their peers as they were about themselves. For the most part, teachers discussed student survey results with each other, both informally and in PLC meetings. They often seemed to know how their colleagues had fared, although in some cases, they were going on what they heard students say rather than conversations with their colleagues. They did not, however, see value in a public presentation of the results.

- *The results went on a building level comparison chart, not by name, but how you rated amongst your peers, and I just thought that was awful because it serves no purpose other than to basically hurt people's feelings. I mean, what is the purpose of this kind of data? If it's to make us better, it's hard to see how this would accomplish that.*

*Suggestions Regarding Student Engagement*

While they thought the information was useful, many teachers (around 20%) suggested not using a student engagement survey as part of the teacher evaluation.

- *I think the student survey would be better used as a reflection piece, but not as a basis for making an evaluation decision to keep or not to keep a teacher.*
- *Because there are so many variables with it, you know, I think it's pretty tricky to tie it to teacher evaluations.*

Some wanted to get more information when given a negative rating, so students could explain why they answered as they did.

- *It would be more useful if students could comment on things . . . You know, if they're giving me a really low score I want to know why.*

Related to this, a teacher wondered if students could be given an opportunity to rate their own effort in a course.

Many teachers suggested developing surveys particular to a school's departments and curricula. At least half of the teachers interviewed had done this in the past, and some continued to use their own surveys this year in addition to the State Model Student Survey.

Another suggestion was to account for the fact that some teachers had large numbers of students responding to a survey, and others (e.g., special education teachers) had very few. This was viewed as potentially unfair, since one negative survey in a group of six would bring down a teacher's student engagement rating much more dramatically than one negative survey out of 70 responses.

<div align="center">

***Overall Comments on the Model***

</div>

All 37 teachers (from all eight interview districts) were asked questions about their overall experience with the Model. The teachers' comments about the benefits and challenges of the Model are presented below.

### *Benefits of the Model*

Twenty-two out of 37 teachers (59%) mentioned that the focus on their instructional practice was a benefit of the Model. They talked about this in a variety of ways: planning, goals/purpose, energy, ideas, reflection, growth/change, engagement, and accountability. Many teachers indicated that the Model asked them to do things they were already doing on their own, but with more "structure and follow-through."

- *I think you should always be trying to hone your craft, and try to do things that are going to engage your students more.*
- *It made me more aware of little things that I could change, you know, with some of my students. Like working in small groups, being more effective. Yeah, I think there were some changes that came out of it that were good.*
- *I think having this as a growing experience rather than say a punishment is a good thing. I mean, students change, the way students learn, so constantly learning for us is best so we can keep up with them.*
- [Without the Model] *I don't think I would have been as focused on my goal because there're always so many things to deal with, so many fires to put out.*
- *It's made me more reflective on "What am I doing? Why am I doing it?" Also, it* [the Model] *created the awareness that people can continuously get better at their trade. We shared that with the kids in our classes and I think overall from both teachers and*

*students there was some growth in the idea that no matter what you do, you can always get better at it.*

Twenty-one out of 37 teachers (57%) brought up increased meaningful contact with their colleagues as a benefit. Of these, the most enthusiastic responses came from districts or schools that worked on the parts of the Model in new or newly energized PLC groups.

- [About working on the Model with PLCs]: *There were some times where you're like: "This is really rewarding and I learned a lot today." How often do you get that with a PD* [professional development] *experience?*
- *I felt that this was the most rewarding year I've had. The students—they've always been the rewarding part—but professionally I felt like for the first time we, as a staff and as a school, just started talking about what it takes and what it means to be a good teacher.*
- *Everybody, I mean everybody, when they talked about positives talked about the PLC. That was the big "Yes!" They loved it. Just being able to talk to each other.*
- *Collegiality: Wonderful. I am in awe of what my colleagues do here. I'm in awe of these people—their passion, their expertise, their knowledge—I have a lot to learn here.*
- *It forces teachers who were . . . kind of disconnected to actually talk with someone about what they are doing, or what was going on in the classrooms. And I think it also gave us all a little bit of the bigger picture.*
- *I connected with the staff more this year than any other year. I think this program helps us be a little more cohesive and learn more from others.*

### *Challenges of the Model*

Twenty-seven teachers (75%) mentioned time when asked what challenges they encountered in working with the Model. Their descriptions ranged from mild to extreme frustration about the amount of time spent on the various parts of the Model. Some teachers were very concerned about what would happen in the future when their district would not have financial support from the state to cover implementation costs..

- *The challenge is the time. How do we, I don't want to say reward, but how do we find some benefit for those who are peer reviewers so they will keep doing it? It takes a lot of time!*
- *All the documentation is a ton. The paperwork is repetitive. There were just too many components and too much paperwork. We want to really streamline and have people not feel burdened by it.*
- *Paperwork, paperwork, paperwork. There's got to be a simpler way to document things.*
- *They* [colleagues] *felt rushed. I felt rushed.*

One third of the teachers interviewed (in six out of the eight districts) talked about the fact they did not understand the "big picture" for much of the year. They hoped things would become

clearer to them in the future, and indicated that the confusion added to their stress. A number of teachers felt there was a failure to communicate the way things should fit together. What might have been the administration's effort to not overwhelm teachers with too much information about the Model may, in fact, have contributed to their feelings of frustration in "not really knowing what piece was coming next."

- *If there was like a year-long calendar posted in the faculty lounge showing the big picture. As simple as that sounds, I think that would have kept it more clear, made things a little less stressful, if we could see the whole process.*
- *It's just getting over the hurdle of understanding what the purpose is of each of the parts. Just like our students, we need to know why are we doing this, and what should it look like.*
- *Just keeping it straight in your head. The difference between what you're doing in the PLC, what your growth plan is, and then the student survey and student growth, whatever it was. Keeping all those balls up in the air was … interesting.*
- *I still have to stop and separate them* [IGDP goal and student learning goals]. *I confuse them easily. The pieces are close enough and I maybe foolishly, maybe not, put them both with the same class. I still have to stop and separate them in my mind.*
- *The entire process has been so confusing. I needed an overview and I don't think I ever got it.*
- *It seems that with all the different pieces and all the different goals some parts were redundant and some parts were just confusing. We're planning to simplify the forms.*
- [In the future] *I think we need time to work on our goals, in our own schools. It's wonderful for show and tell* [PD] *but I need somebody to sit down right next to me and say, "Okay, you put it in this file and then you can do this, and here's the access code," etc.*

### *Additional Comments*

Teachers had many miscellaneous comments about difficulties with the Model, from their introduction to it, to the tension of being called out of their classrooms. They were hopeful that it would go better in the future, and the majority were certainly not averse to being evaluated.

- *Just the initial training was sort of hellish. That's not really the Model, but it was quite tedious. But the Model, it's probably really over-engineered. It probably needs to be simplified, so it's doable.*
- *Initially, you know, people became anxious. It needed to be introduced a lot better.*
- *It's really hard to be out of your classroom. That was my biggest complaint.*
- *It wasn't perfect, but it wasn't horrible, either.*
- *I think that we need to be held accountable for what we do, there's no doubt about it.*
- *We had a really good year with this, which I thought should have been a catastrophe, but it wasn't.*
- *The authenticity of the process drove it forward, rather than fear of or approval from the principal.*

*Summary of the Teacher Interview Findings*

- The majority of teacher interviewees (70%) mentioned benefits of increased collaboration and stronger relationships among faculty resulting from points of contact with peer reviewers. However, teachers found points of contact with peer reviewers to be terribly time-consuming when it came to documentation.

- Those peer reviewers who did classroom observations voiced the challenge of leaving their own students even if a substitute teacher was offered for that time period. Although they commented on how much they learned from each other when they were able to visit one another's classrooms, they also requested clarification regarding their role. Was the function of the peer reviewer that of a mentor or an evaluator?

- Another negative aspect of points of contact with peer reviewers reported by some teachers was that they did not feel comfortable being honest with their peers about what they saw in the observations. In situations where peer reviewers were reluctant to give honest feedback, collegial relationships did not flourish despite increased contact between teachers.

- Teachers talked about varying levels of quality in the points of contact with summative evaluators. Key factors included the amount of time allotted to the pre- and post-observation conferences and the specificity of the feedback. Teachers appreciated when the points of contact were focused and specific, and several attributed this to the use of the IGDP and the rubric.

- The majority (75%) of teachers found the IGDP to be very useful, with the remaining teachers expressing more neutral feelings about the process. The most common reason expressed about why the IGDP worked well for teachers was that it kept them focused. Reasons for less enthusiastic responses included dissatisfaction with their professional goal, such as setting a narrow goal that was too easily achieved or one turned out to be difficult to achieve.

- Overall, teachers found the self-assessment process useful, although some thought it was overly time-consuming.

- Over 45% of the participating teachers used an internal measure to set their class learning goal, whereas just over one third used an external measure (e.g., MCA or BAS). Most of the remaining teachers (15%) also used internal measures, but they sought "improvement" rather than specific quantified results.

- Forty percent (40%) of the classroom teachers interviewed said SLGs were useful. The teachers who found SLGs useful talked about the way the goals increased their focus, planning, and reflection.

- Specialist teachers found the SLGs to be somewhat challenging because of complications due to subject matter, larger numbers of students, and fewer meetings with all students in any given week.

- Nearly all special education teachers interviewed found the development of SLGs to be redundant, given the fact their students already have individual learning goals in their Individualized Educational Plans (IEPs).

- Roughly half of the teachers interviewed felt their students had met the class learning goal they had established, although only 18% had supporting data. Those without supporting data who felt the class learning goal had been achieved spoke about being satisfied with the progress of their students.

- Two thirds of the teachers interviewed said they had set targeted group goals, while one third did not designate a targeted group. Twenty percent of teachers talked about the fact they did not have an adequate number of students to warrant both class learning goals and targeted learning goals. Teachers in most districts were neutral or unenthused about drawing distinctions between groups of students.

- Teachers who commented on problems with the SLGs (45% of those interviewed) found them unwieldy for a number of reasons, particularly noting a need for more guidance and training on how to measure growth and then use that data to change instructional practice.

- Of the teachers interviewed, 40% were positive about the student survey experience, and a few more (45%) thought a similar survey should be used in the future. Positive responses tended to touch upon the importance of hearing students' voices.

- Forty-two percent of teachers expressed concerns about validity of the survey data, citing the age of their students as potentially affecting their ability to comprehend the survey items or value the survey's purpose. An additional area of concern was the way the State Model Student Survey data compared teachers against one another. This seemed unnecessary to many of the teachers, and at least one teacher at every interview district expressed concern for their colleagues in the event of low student engagement results.

- Perceptions about the Minnesota Standards for Teacher Practice rubric were decidedly split between those who found it useful and those who did not. Those who spoke positively mentioned that it served as a structure on which to build and grow and that it would provide stability over time. Those who spoke negatively believed that the rubric limited thinking and led to canned ways of reflecting about their work.

- Overall, 59% of the teachers mentioned that a benefit of the Model was the focus on their instructional practice. Many teachers indicated that the Model asked them to do things they were already doing on their own, but with more structure and follow-through. However, 75% mentioned the amount of time it required as a challenge. In addition, a number of teachers were very concerned about what would happen in the future when their district would not have financial support from the state to cover implementation costs.

# ♦ SUMMATIVE EVALUATOR INTERVIEWS

## Methodology

### *Interview Procedure*

During May and June, 2014, interviews were conducted with summative evaluators by phone, each lasting about 30-45 minutes. Interviews were audiotaped, transcribed, and analyzed for common and unique themes.

### *Interview Protocol*

The summative evaluator interview protocol included questions regarding teacher practice, student learning goals, and student engagement. A copy of the interview protocol is included in Appendix I.

### *Participants*

Eleven school administrators in eight school districts were individually interviewed. The 11 administrators had also been interviewed by CAREI in the fall and had been selected based on the extent of their participation in the Pilot. Six of the eight districts had agreed to implement all components of the Model (i.e., teacher practice, student learning and achievement, and student engagement) over the course of the Pilot year. Two of the districts opted to implement only one or two components of the Model. All eight districts had one or more summative evaluators, since all of the components they implemented involved the assessment of teachers as part of their participation. All 11 interviewees had a formal role in the teacher evaluation system as a summative evaluator because they were all school principals. According to the licensure of principals, they have the role, authority, and obligation to determine the instructional effectiveness of individual teachers. Principals are required to observe and evaluate all teachers, with tenured teachers having an evaluation cycle once every three years, and probationary teachers having on-going evaluations each year for three years before being granted tenure. In order to protect anonymity, the components discussed by individual summative evaluators or across districts are not identified.

## Findings of the Summative Evaluator Interviews

Findings from the summative evaluator interviews are reported by the components of the Model: Teacher practice, student learning goals (a sub-component of the student learning and achievement component), and student engagement. The summative evaluators' responses regarding the overall usefulness and effectiveness of the Model are provided at the end of this section.

### *Teacher Practice*

The teacher practice component has three required parts, including: a) the development of an Individual Growth and Development Plan (IGDP); b) points of contact between teachers, peer reviewers, and summative evaluators; and c) use of the *Performance Standards for Teacher Practice Rubric*. A fourth part, the development of a teacher portfolio, is optional.

Within the first month of each school year, all staff on a teacher contract develop an IGDP in which they identify one or more areas for professional growth that year and the methods or means that will be used to reach their goals. The summative evaluator is expected to read and initially comment on each teacher's IGDP in the fall, and then use it as a guide throughout the year in conversations with the teacher about his/her professional growth.

The points of contact are a way to keep track of the summative evaluators' and peer reviewers' professional conversations with and observations of a teacher. The summative evaluator is to carry out two points of contact with each teacher every year. Each teacher also has at least three points of contact with a peer reviewer every year. Altogether, five interactive events will be recorded for each teacher every year.

The teacher practice component also includes the use of a rubric developed for the Pilot program (i.e., Performance Standards for Teacher Practice Rubric) which is a self-assessment. Finally, the component incorporates the optional development of a teacher portfolio.

The summative evaluator interview included questions about the usefulness and challenges of each of the features of the teacher practice component. Key teacher practice component findings and a summary of findings for each feature of the teacher practice component are presented below.

### Key Findings

Overall, the teacher practice component was viewed positively by all summative evaluators. More than half of the summative evaluators noted that this was the best part of the Pilot. Many respondents referred to the value of having a structure in the form of the IGDP to facilitate conversations with teachers. The IGDP enabled the teachers to choose their own areas for personal and professional growth, and thus allowed their principals to engage in meaningful, reflective conversations with each of them about their work.

Specifically, the setting of personal growth goals by each teacher *"led to discussions that led to conclusions that should lead to change,"* observed one summative evaluator. Another noted that *"setting personal goals is changing the way that the teachers are thinking."*

     ***Points of contact***.  The Model was designed to have five points of contact documented annually for each teacher. All summative evaluators noted a range of benefits of using the points of contact as part of the teacher evaluation Model. For example, the points of contact:

- *Helped to organize the PLCs*
- *Provided good evidence of professional conversations and growth*
- *Provided an opportunity for staff to communicate with each other*
- *Helped us to have a lot of structure*
- *Promoted collaboration among the teachers*
- *Were the best part of the pilot*

The points of contact also were seen as a useful tool in the peer review feature of the teacher practice component. Nearly every summative evaluator noted that teachers appreciated having points of contact because, as one principal stated, *"The recorded points of contact validates and communicates to administration the conversations [*between teachers] *that were already going on."* Another principal stated that the peer review feature encouraged *"more professional conversations than just* [being about] *teacher evaluation."*

Negative perceptions of points of contact were primarily centered on the paperwork needed to document their occurrence and provided evidence of the extraordinary number of documented events some of the summative evaluators had to complete. The following comments exemplify the principals' concern about the time it took to document or conduct points of contact:

- *In some districts, like mine, where one administrator is in charge of a significant number of staff –for example, I had 60 teachers and 17 of them were non-tenured—my documenting all of those points of contact is not sustainable.*
- *It is unreasonable to expect two points of contact to be done by each principal for each teacher.*

Interestingly, only one of the summative evaluators agreed with the suggestion that the information collected by the points of contact led to planning of ongoing professional development in his/her building. This person went on to explain that his/her district *"did professional development on how to do a peer review, including having a session on inter-rater reliability training."*

The comment most frequently expressed by the summative evaluators about points of contact was their desire to do more classroom observations than observations of PLC or grade-level meetings because classroom observations provided more meaningful information. As one principal said, *"Observations are more meaningful than meetings."* Another who valued classroom observations said, *"There was not enough time to do full class visits, so I did more quick drop-ins."* Still, another said, *"I was unable to get into the classrooms as much as I would have liked."*

> ***Performance Standards for Teacher Practice Rubric***.  Another feature of the teacher practice component is the rubric developed for the Model by the Work Group Task Force and MDE leadership. When specifically asked about whether or not it was a useful tool in the assessment of the teachers, all of the summative evaluators agreed the rubric and its link to standards of instruction were helpful. Overall, the rubric and related standards for teacher practice were viewed as a feature of the evaluation process that should be continued. The following comments reflect what was good about the rubric and what could be improved.

> Positive comments:

- *It was useful for guiding conversations.*
- *It had a good emphasis on growth and support.*
- *The* [score of] *four allowed for recognition of truly exemplary teaching.*
- *It has good components.*
- *Teachers were able to use it as a self-assessment.*

Areas needing improvement:

- *Some teachers took a lot of time* [to complete it], *while others did not. Some under-rated themselves, while others did the opposite.*
- *It was useful, but still needs on-going training to use it well.*
- *Some of the language in the rubric is vague.*
- *It took too much time for the assessors.*
- *Scoring is not very useful—the rubric is more useful for starting the conversation.*

One principal explained the difficulties by expressing a reflective concern that the confusion over the rubric could be addressed if he/she had had more discussions with the teachers about the language in it, stating, *"I had a batch of teachers who did not have an understanding of the process.…You know, understanding the rating and what the rubric really meant…It's probably a lack of my leadership, not having this person understand…what the words within each rating meant. …There were a couple of teachers who were not prepared at all* [to do the rubric]*…they did it because they had to."*

**Teacher portfolio**.  Of the eight districts included in the interviews of summative evaluators, only two had a few teachers who chose to develop the optional portfolio. For the teachers in those two districts, the summative evaluators reported that they and their teachers *"struggled with the appropriate format, length, and detail to include in the individual portfolios"* which are a combined record of the accumulated evidence of teacher practice. Some of the teachers also included short video segments of their classroom instruction as part of their portfolios. The summative evaluators who reviewed the portfolios were not exactly sure what to do with the greater amount of documentation for those teachers as compared with the vast majority of teachers who did not compile a portfolio.

### *Student Learning Goals (SLG)*

Nearly all of the summative evaluators who spoke in their interviews about their teachers setting SLGs indicated that this particular component was moderately successful at best.

### *SLG Training*

Several of the principals mentioned the training that was provided to their staff about how to set measureable goals to produce useable data occurred a bit too late in the process. Thus, some districts' teachers were crafting SLGs early in the school year when they had not yet been trained on what was needed to fulfill this expectation. As one principal recounted, *"At the end of September we had the training on setting student learning goals, but the teachers had already submitted them to me. Many of my teachers said after the training, 'If I had known that, I would have done them differently'."*

*Successes and Areas Needing Improvement*

According to the interviewees, several districts found that using their professional learning communities (PLCs) to facilitate the goal-setting process worked well, and they would do it that way again. Understanding what worked and what did not is clearly reflected in the following range of comments:

- *I realized that teachers found it difficult to shift from assessing assignments to assessing* [student] *growth.*
- *For the most part, our teachers set appropriate goals. However, some were afraid to set a lofty goal [*and then fail*], while others set goals that were too high and unrealistic.*
- *Some of my teachers set super low goals because they were scared…since how their students did turns into a rating for them. …It is a little uncomfortable when* [I have to explain to them that] *their rating score is not a 4, and* [I have to say] *that it is because their goal wasn't rigorous enough to begin with.*
- *At our school, some teachers found it* [the setting of SLGs] *unfair because they were required to use MCA scores, while others* [not teaching reading or math] *could pick whatever assessment and goal target they wanted.*
- *Our teachers' goals focused too much on knowledge and recall. As a result, this did not lead to changes in teacher practice.*
- *We need to have more conversations about meeting standards versus showing improvement.*
- *I had a lot of teachers reflect that they would do it a little differently next year. They might choose a different assessment...because a lot of it is about the assessment itself. …They are reflecting on the availability of getting good data.*

Finally, several summative evaluators explained why and how the setting of SLGs was an important component of the Pilot. As one noted, *"I think the strength of the pilot was that classroom teachers could set their student learning goals based on their own assessments, as long as they were approved* [by me]*."* Another principal saw the value in setting goals and assessing student learning because, *"for me, it differentiated strong from weaker teachers."* And a third principal said, *"Having the data provided tools for talking about struggles."*

*SLGs Set by Non-Classroom Teachers*

Issues related to how specialists and non-classroom teachers struggled with setting student learning goals were mentioned by all respondents. Most noted that art, music, media, and physical education teachers were able to set measureable goals primarily because they had a curriculum and content standards to use in guiding what to assess. What worked in these situations was to have peers in the same subject area work together to set goals and provide feedback on them to each other. When that was not possible, one principal said, as an example, *"Our arts teachers had to figure out how to communicate goals and assessments to non-arts teachers."*

Other professionals who are on teaching contracts, but who do not have a classroom, such as psychologists, nurses, and counselors, had greater difficulty than classroom teachers in setting

"classroom" goals but were able to set "targeted group" goals. The summative evaluators in many schools thought that the ease of setting goals by special education teachers was also mixed. They reported that some special education teachers found setting SLGs very easy, while others felt it was redundant because learning goals are already included in each student's Individual Education Program (IEP). Principals mentioned that some of the same issues for special education teachers affected Title I teachers as well, saying for instance, *"My Title I staff worked with so few kids—that was a bit challenging. ...We had* [targeted group] *goal sets from fall that we measured in the spring, and the kids had already been out of their* [targeted] *groups since December."*

### *Student Engagement*

This component had the widest range of opinions about its usefulness of any of the components in the Model. Three of 10 respondents who spoke about this feature felt very positive about all aspects of it; the others had mixed views. The positive reactions from all principals centered on the concept of the value of including the student's voice in assessing a teacher's performance. One summative evaluator expressed its value with the following comment, *"If I was a teacher, it [the student engagement survey] would cause me to reflect every day on...what the students need to engage to learn....It gives a teacher a rubric of self-assessment on the work they do with kids every day."* Another principal noted, *"It was validating for teachers, and it gave students a chance to have their voice heard as well."*

The comments from the summative evaluators about the value of the student engagement survey also tended to cluster around its usefulness to complete the picture of a teacher's performance. Similar positive remarks include:

- *This* [the student engagement survey] *is one of the most important parts.*
- *It was good to have a school-wide survey...The students enjoyed getting the chance to share feedback on teachers, and for the most part, they feel very engaged with their teachers.*
- *I was able to use the engagement results effectively in the summative conversations with my teachers—especially for individual professional development needs.*
- *I was careful to keep in mind the limitations of the survey—that it is just a snapshot. I need to watch for patterns over time.*
- *I used the results in tandem with classroom observations.*
- *I also invited teachers to add their own measures or documents on student engagement.*
- *I think the added component of having the student survey was huge. The teachers, a lot of them,* [this year] *got feedback from their kids differently, so to have that survey administered across the school was pretty big.*

The negative comments about the student engagement survey centered around two main themes—the difficulty in scheduling the administration of the survey, and concerns about the wording of questions, especially for the younger students.

- *It was difficult to navigate the administration when there was more than one teacher in a class.*

- *Timing was bad – it was administered at the beginning of a new semester* [when the students did not know the teacher very well].
- *Only three class hours were surveyed, so teachers had variable samples. The percentages were often very misleading.*
- *The wording of the questions was confusing.*
- *My younger students couldn't understand the questions.*
- *Some of the teachers didn't like it because they thought the student scores were unreliable.* [When the results came in], *they saw some of the best teachers in the school get the lowest scores. Also, some teachers watched kids bubbling random answers on their survey forms.*

In the end, nearly all summative evaluators mentioned during their interview that they cautioned their teachers *"not to take it personally."* They noted some teachers still had confusion about why the student engagement survey was used or useful: *"I think it just depends on how much energy they invested into their own results and learning about themselves."* Several principals commented they had good training from MDE staff (Tyler Livingston and Greg Keith) on helping the teachers interpret the results, and yet also said, *"I probably could use a little more training to better interpret those results...but also this is just our first year to do this."*

### *Teacher Performance Rating*

Five of the 11 summative evaluators gave each teacher a composite rating on their performance at the end of the year, and the other six did not. Those who did not provide a rating to teachers either ran out of time at the end of the year or had their teachers do a self-rating. *"I had them do a self-rate and then I commented and reinforced that. For the most part, they were dead-on. At the end of the day, I wanted this to be a reflective process...that's powerful, I think."* Another principal commented, *"I said to each teacher, 'What does the data tell you? You tell me where you need to be rated.' I put the ball in their court. For this year, it was a meaningless rating, but good to go through the process."*

All but one of the summative evaluators who did provide each teacher with a teacher performance rating score believed that having a numerical score did not have much value. Their comments included:

- *Yes, I gave each teacher a score, but I am not sure if it is meaningful.*
- *Determining a rating score was more work than it provided useful information.*
- *Preparing for a summative conversation would take me two hours per teacher...even though I had an Excel spreadsheet to crunch the numbers. It's a balance between subjective data and looking at the evidence, but at the end of the day, you are making a call. Do I think that the number that I would spit out at them was always completely accurate in terms of its being truly reflective of them, No, I don't.*
- *I think conversations without numbers would be at least as meaningful.*
- *It's one of those things that, in the end, it doesn't matter because the summative evaluator can disregard or change the rating for a lot of very good reasons...such as having a bunch of new students come in. I don't think that it really measured what it was intended to measure.*

The one principal who thought that determining a rating for each teacher did have value noted, *"I met with all of my teachers one-on-one at the end of the year. It took a lot of time—I had 45 that I did. It was pretty tough at the end of the year to meet that goal, but I got a lot out of it, and I think a lot of my teachers did too."*

### *Overall Perceptions of the Model*

We asked the summative evaluators to tell us about the Model overall and the effect that it may have had on school culture among the staff, its effect on students, and changes that they might suggest for the future.

### *School Culture*

Nine of the 11 summative evaluators perceived positive value for the changes in school culture as a result of participating in the Pilot program for teacher evaluation. Overall, two main themes emerged from the comments—the first being a sense of professionalism among the teachers. As one principal stated, *"It was positive…It gave people a sense of pride because we have MDE coming here, and we are learning these things, and MDE is learning from us."* The second prominent theme was the sense that the teacher evaluation system led to perceptions of greater accountability among the teachers. One principal used the term of *"healthy panic"* to describe what was happening in his/her school:

- *Toward the beginning of the year when we were first setting goals, and people were looking at kids' scores "meeting expectations" or being "below mastery level," there was what I called "healthy panic" where* [after teachers looked at student scores on an assessment] *people said, "We need to do something, we have to do something about this."… I'd hear people in the lounge, or when I'd walk through the halls, talking about their goals and their growth plan, and when they were going to meet…. It just kind of forced everybody to be accountable for that stuff.*

Other positive sentiments expressed by the summative evaluators include:

- *I think I maybe did a better job than I have in the past as a principal having good back and forth conversations between the teacher and me, where it wasn't just me going into a classroom, gathering information, and then telling them what I thought it meant.* [In this Model] *it was a lot more. The teachers have power to submit certain pieces of evidence to me and contribute to the conversation in a meaningful way. I liked that.*
- *This new system "shook their nests" and it created new thoughts for us to gather around.*
- *I think it is all positive. We are shifting to an assessment, goal-setting, accountability culture—something that is all new for this school.*
- *Participation in the pilot brought staff together, with some of it related to resistance.* [We now have] *a culture of teacher evaluation, when previously there had been none.*
- *Accountability had a positive effect on the whole school.*
- *We went from having little or nothing* [for teacher evaluation] *to one of the more rigorous processes that there is.*

- *I made this as serious as I could without making it scary.*

### Effect on Students

Only 4 of the 11 summative evaluators believed the implementation of the teacher evaluation system had any noticeable effect on students. The one negative view was a perception that some students were *"rebelling because the school now had more expectations on how they ought to behave."* The positive benefits were centered on the global view that instruction improves when teachers set goals. One principal simply stated, *"I think it"* [participating in the Model Pilot] *positively affected our students."* A response from one of the summative evaluators summed it up by saying, *"Yeah, of course. I think it definitely benefits the kids with their teacher setting goals, and improving themselves and their instruction, and letting data kind of drive what they're doing in their classroom. It had a great effect."*

One principal just replied *"no"* when asked whether the evaluation system had an effect on students. The other sentiment expressed by two other principals was that *"students wouldn't recognize a change."*

### Final Thoughts and Observations About the Model

At the end of each interview the summative evaluators were asked about what they thought about the process overall, and if they had any plans for changes that they might be making next year in their local teacher evaluation system. Ten of the 11 persons interviewed used positive language to describe how their district experienced the Model. This is what six of them had to say about their implementation experiences:

- *I learned a lot about my teachers individually, and I learned a lot about my grade levels collectively…and I learned how my teachers plan. … It gave me a pretty good grasp of my building, so all of the components I feel helped me as a leader.*
- *After looking at the files, some of the teachers here, believe it or not, had not been evaluated in 20 years…So what this* [pilot] *did was it brought this set of teachers up to speed very quickly on what evaluation is and what it's supposed to be.*
- *The best part was the professional conversations.*
- *I emphasized to staff that this is about working together and not about punishing anyone.*
- *Having this* [the state-wide evaluation system] *encourages the top teachers to improve as well.*
- *The training and support that we received from Tyler and Greg was awesome.*

The one person who was hesitant to give the process a *"thumbs up,"* said that his/her views were *"mixed"* about the entire experience. Although the vast majority had good things to say, as seen in the comments made for specific aspects reported in this summary, each person also tempered his/her response by expressing some form of concern that will likely lead to changes in what is done by each district next year.

- *The whole issue, I think, for small districts is going to be cost. We could've never done what we did* [this past year] *without that grant….I think one of things that have to be*

> *brought out to legislators is that when you mandate things, you have to give us the money to do that. Because when it comes down to hiring subs for teachers to observe other teachers and for leadership teams to meet, and such, or do a roof repair, I'm going to repair the roof. And I'm not trying to sound like sour grapes. I'm saying the intent of the legislation* [for teacher evaluation] *is good, but* [to continue this work] *we need the funds.*

- *If there is not going to be any state oversight, it will tend to go back to not having the value that it [*teacher mandated evaluation*] should have. I think the idea is really good, but if there is not state oversight of any kind, and it's just left to the districts, well it might have no teeth...As soon as we told our teachers that there is no state oversight, it was like, this big sigh, "OK, good."*

The most prominent negative theme that was spontaneously mentioned by every summative evaluator was the workload and the amount of paperwork involved in the evaluation Pilot.

- *Administratively, it was—and I hate to use the word—overwhelming. It really was.*
- *It was overwhelming.*
- *I found the workload at the end of the year to be overwhelming.*
- *It was a pretty tough end of the year, but I got a lot out of it.*
- *The problem that we have in our schools, and I believe in the value of all of the changes, is the rigor of all of the new things, and that everything we're asking people* [both teachers and school leaders] *to do is adding to their workload.*

### *Summary of the Summative Evaluator Interview Findings*

Eleven summative evaluators in eight school districts were individually interviewed. All interviewees had a formal role in the teacher evaluation system as summative evaluators because they were school principals, and the components they implemented involved the assessment of teachers as part of their experiences with the Model.

The **teacher practice component** was viewed positively by all summative evaluators, and more than half of the summative evaluators thought that this was the best part of the Pilot. Specifically:

- The Individual Growth and Development Plan (IGDP) that each teacher developed was seen as valuable because it enabled the teachers to choose their own areas for personal and professional growth and it served to facilitate professional conversations between a teacher and his/her principal.
- The points of contact were seen as a useful tool because they validated professional conversations, and they shifted the evaluation system away from being only a compliance and regulatory system to one of professional growth.
- Negative perceptions of points of contact were primarily centered on the paperwork needed to document their occurrence. The comment most frequently expressed by the summative evaluators about points of contact was their desire to do more classroom observations than observations of PLC or grade-level meetings because classroom observations provided more meaningful information.

- The rubric for performance standards for teacher practice was viewed as a useful feature of the evaluation process that should be continued because it has an emphasis on growth and support. A few concerns were expressed, however, by the summative evaluators that mostly related to vague language in some of the standards and the lack of time and training for them, as leaders, to effectively use the rubric to stimulate professional growth conversations.
- Only two of the eight districts had teachers who chose to complete the optional teacher portfolio feature of the teacher practice component. Summative evaluators were unsure of how to use that greater amount of data in developing a rating score for those teachers.

The implementation of the **student learning goals** element of the student learning and achievement component was seen by all summative evaluators as only being moderately successfully at best.

- The positive feelings about student learning goals concerned the setting of measureable learning targets for students. Principals thought the data from the learning goals provided concrete information and a good conversational focus for where teachers could improve.
- The negative perceptions were centered not on the concept of setting learning targets, but on the need for more training for some teachers to know how to develop and set measureable goals.
- Also, issues related to how specialists and non-classroom teachers struggled with setting measureable student learning goals were mentioned by all respondents.

Of all the components in the model, the **student engagement component** had the widest range of opinions about its usefulness, especially the usefulness of the student engagement survey. Three of 10 respondents who spoke about the student survey felt very positive about all aspects of it; the others had mixed views.

- The positive comments from the summative evaluators about its value tended to cluster around its usefulness to complete the picture of a teacher's performance.
- The negative comments about the student engagement survey centered around two main themes—the difficulty in scheduling the administration of the survey, and concerns about the wording of the questions, especially for the younger students.

The **teacher performance rating**, which was intended to be a composite score based on the outcomes for the Model's three main components for each teacher, given by summative evaluators at the end of the school year, was only partially achieved. Five of the 11 summative evaluators gave each of his/her teachers a composite rating on their performance at the end of the year, and the other six did not.

- Four of the five summative evaluators who did provide each teacher with a teacher performance rating score believed that having a numerical score did not have much value because the score was not perceived to measure what it was intended to measure. One principal thought that both he/she and the teachers benefitted from having a summary score.

- Those who did not provide a rating to teachers either ran out of time at the end of the year or had their teachers do a self-rating as a self-reflection exercise.

The summative evaluators were asked about their **overall perceptions of the Model.** Nearly all perceived positive value for the changes in school culture. Two main themes emerge from the comments:

- A sense of professionalism among the teachers was created by having school-wide participation in the teacher evaluation Pilot.
- The teacher evaluation system led to perceptions of greater accountability among the teachers.

Every summative evaluator had praise for the training and support provided by MDE. Ten of the 11 persons interviewed used positive language to describe how their district experienced the Model. All commented that they had learned an enormous amount by participating in the year-long Pilot, despite the overwhelming amount of paperwork involved in it. Furthermore, for the next school year, most districts were planning to take many of the features of the Model and use them in their district's creation of a modified version of a teacher evaluation system.

# ♦ MDE STAFF INTERVIEWS

## Methodology

### *Interview Procedure*

Individual interviews with MDE staff responsible for the development and implementation of the Pilot were conducted in July and August 2014. The interviews were conducted in person and were audio-taped, transcribed, and analyzed for common and unique themes.

### *Interview Protocol*

The MDE staff interview protocol included questions regarding areas of success and challenges in the Pilot implementation, possible revisions to the Model, role of MDE in the future, availability of training, and lessons learned during the Pilot year. A copy of the interview protocol is included in Appendix K.

### *Participants*

Two MDE staff members were interviewed. They were Tyler Livingston, Educator Effectiveness Supervisor; and Greg Keith, Director of the Division of School Support.

## Findings of the MDE Staff Interviews

The responses from both interviewees are combined in the findings summarized below. Initials are used to attribute quotes to either Tyler Livingston (TL) or to Greg Keith (GK).

### *Successful Aspects of the Pilot Model Implementation*

The interviewees described success as occurring on two levels. First, the overall Model appears to have successfully encouraged teacher buy-in to an evaluation process. Second, the Pilot presented the opportunity to learn which aspects of the Model were most effectively implemented and why.

The willingness of districts to participate in one or more components of the Model and their cooperation in carrying out required activities was cited as being critical to the success of the Pilot. The collaborative nature of the relationship between MDE and the Pilot districts was considered key.

Additionally, the Pilot districts were effective in sharing feedback with teachers and offering guidance on how teachers might reflect on and utilize the feedback to improve their instructional practice. Teachers and summative evaluators frequently referred to the sense of partnership that developed between them when they met to discuss an observation or to go over plans, goals, or survey results. The process of teachers and summative evaluators working together for a common purpose was perceived to have a very positive impact on teacher buy-in. The

interviewees considered points of contact and feedback to be the most successful individual elements of the Model.

### *Comments Regarding Successful Aspects of the Pilot Model Implementation*

- *We've had great partners in the work. They've all, I think, operated with the wish to see the pilot succeed and to have good practices for the leadership and the teachers. I think that has been a strength. They've always been willing to work with us and collaborate with us. (TL)*
- *Probably the feedback* [to teachers] *was most successful. Through the course of the year when cluster observations or points of contact were conducted, I think that probably was the best part... Maybe not the number of points of contact, as I think there was some struggle with that... I think just in terms of teachers reflecting on their practice, I think that was pretty successful. I get the feeling that that went pretty close to what we were expecting. (GK)*

### *Future Modifications to the Model*

The implementation of the Pilot was a learning experience that helped identify aspects of the Model where modifications would be beneficial. An idea that surfaced multiple times during the MDE interviews was having a phased implementation strategy, rather than rolling out all pieces of the Model at once. Both summative evaluators and teachers alike indicated they would have appreciated an initial introductory overview of the entire Model followed by step-by-step training for each individual component. A phased strategy such as this might have reduced confusion and feelings of being overwhelmed by the Model's expectations.

Possible revisions to specific aspects of the Model were also identified by the interviewees. For example, it was mentioned that the number of points of contact might be reduced. It was also suggested that the peer review process be analyzed to identify the training and ongoing coaching that teachers need in order to become effective as peer reviewers. The peer review process was considered to be one of the most valuable elements of the Model, and improved peer reviewer training could make it even more so. One surprising outcome was the difficulty that teachers had with the creation and measurement of student learning goals. It was suggested that increasing training on how to set, assess, and reflect on data from student learning goals would likely eliminate much of the difficulty.

A common theme in summative evaluator feedback was that implementation of the Model's multiple elements was a significant time commitment. When asked about this, both MDE interviewees noted that processes generally take longer when an initiative is first implemented than later on after it has become more routine. After the Model becomes familiar and summative evaluators and teachers have acquired experience with it, it is likely that the components will take less time and effort. Therefore, the Pilot year might not provide good estimates of time required in the long term. It is also likely that summative evaluators will identify ways that the time commitment can be reduced without affecting the quality of the evaluation process. For example, instead of performing full classroom observations for each point of contact, summative evaluators might realize they could shorten some points of contact to a walk-through targeting specific instructional practices that would then be followed by a brief written feedback report for the teacher.

The respondents emphasized the importance of learning from failure in order to implement something well. If a summative evaluator found it difficult to perform the required (or recommended) number of observations, a possible problem-solving approach, for example, might be to have the superintendent work with the principal to remove things from his or her plate in order to make it work.

The interviewees also noted that the Model may require summative evaluators (principals) to reconceive their roles as instructional leaders. Summative evaluators' buy-in might increase when they begin to see one of their major roles as promoting teacher growth in support of enhanced student achievement.

### *Comments Regarding Future Modifications to the Model*

- *I believe that Greg and I share the opinion that peer review is very valuable, that the time peers spend with each other to provide feedback, to brainstorm, to review student data, and to share ideas is among the strongest attributes of this work and is a key lever to improve teacher performance and student outcomes. At the same time, I think we would both also acknowledge that teachers need training and ongoing coaching to become more effective as peer reviewers, and that our training during the pilot was insufficient to the need. (TL)*
- *…if I were a school leader, I would say, "You know what? I just couldn't do it because I had these things on my plate. Let's look at our system." And I would hope my superintendent would allow me to be so bold. Okay, "We look at our system to allow me to be the instructional leader that I want to be through this, so what can be taken off my plate, if you will, so that I can do these other activities?"(TL)*
- *They* [districts] *are going to have to rethink about what the principal does. That's plain and simple. Because there is not going to be any easy way to implement what's in that statute. And even if they don't use the state model, they are still going to face the same dilemma. (GK)*
- *I think in some ways principals believed those points of contact to be more than they had to be…They didn't have to be full observations. A principal can do walk-throughs and then give the teacher a written summary of that. (GK)*

### *Implementation of Teacher Evaluation Across the State*

Both respondents estimated 1% of districts are adopting the state Model in its entirety, but they estimated that roughly 65% to 70% of districts will "take something" from the state Model. The interviewees emphasized it was always expected that many districts would not fully adopt the state's Model, but instead would learn from it and adopt the pieces that worked well for them. The intention has been for districts build a model of their own. In order for this to happen, MDE is going to implement communities of practice in which networking, problem-solving, and sharing will occur across districts. MDE is also planning to publish and disseminate materials on promising practices and to work with Education Minnesota and other partner organizations to build a repository of best practices.

It was noted that practices worthy of replication are being used by districts who were not in the Pilot. Those practices need to be disseminated. For example, one of the interviewees mentioned Stillwater's use of inter-rater calibration for teacher observations, which has led to increased consistency and fairness. Teachers' positive perceptions of observations is integral to their buy-in and subsequent use of the feedback.

A variety of training opportunities will be available in the coming year that are designed to meet the needs of districts who adopt the full state Model as well as districts who only adopt parts of the Model. MDE anticipates training will take the form of single-day sessions and webinars on specific processes. For example, a webinar on student learning goals may be attended by several districts who have decided to incorporate the student learning goals component into their own Models. Training on individual components of the Model are being conceptualized toolkit trainings. The hope is that these training sessions eventually will be led by regional centers in the field rather than at the MDE main office. In addition to MDE-provided training, training sessions will be offered by vendors, consultants, and regional organizations like the Service Cooperatives as well as internal district sources. The variety of training resources will allow districts to explore multiple possibilities when creating their own models. An aspect of training that may be emphasized is a focus on areas such as data literacy accompanied by a deep understanding of the Model's overarching goals that will enable leaders to carry out effective evaluation systems.

### *Comments Regarding Implementation of Teacher Evaluation Across the State*

- *Well, we are doing some communities of practice. There's going to be nine communities of practice around the state. And our hope would be that there is some networking and problem solving and sharing in those. One of the things that we proposed that we didn't get funding for was best practices conference where they could share. For principals, we've actually secured the Bush funded FHI 360 to actually publish some promising practices. And then there's the school board association and Ed MN, and we've worked with them a little bit. They are building a repository where districts can post best practices. (GK)*
- *I think State Model training will look two ways. One will be for those who are using the entire system or a close variation of the entire system but then we would build this in, alongside of the version 2.0 Communication, which would be continued training... Saying, "Okay, we've learned this about our student learning goals process. Come to a single day training or participate in this webinar on our newest revised version of the student learning goals process, so that you can improve your local systems, those of you who borrowed something from this." (TL)*
- *I am currently working with six organizations. Five of them are charter schools and one is an [Ed] District. They have said, "Yes, I'm interested in training." Not just training but some ongoing implementation support similar to how I supported the Pilot. So how about a monthly webinar and a spring training for evaluators and things of that sort? The difference for them would be that this is initial training. I'll be inviting implementation teams in for a longer set of time, both training them in the model but also training them as trainers to bring it back to their local sites, because what's unsustainable is to have MDE staff go to every single district's workshop day and lead all teachers. (TL)*
- *One of our next year toolkits or plans is to develop a Minnesota planning guide that is a document a user could pick up, with more training and implementation support around*

*student survey measures, whether they're using this survey we developed for the pilot or others. (TL)*

### Monitoring Compliance with Teacher Evaluation

MDE does not have statutory authority to collect implementation plans. MDE is, however, collecting letters of assurance from districts. Both MDE interviewees believed that teachers and other stakeholders would take action if districts did not have an evaluation model in place that met the legal requirements. MDE intends to provide resources to non-compliant districts in the areas where their evaluation systems needed to change. Additionally, the spotlight on the Pilot has made districts state-wide aware of the statutory requirements and the necessity of implementing an evaluation model. It is believed that the majority of districts will do their best to comply, and that inevitable misconceptions will be ironed out as time goes on. In addition, it seems that teachers are buying into the Model and want the opportunity to get feedback and further develop their skills as educators.

### *Comments Regarding Monitoring Compliance with Teacher Evaluation*

- *We are doing an assurance process…In the fall…the district needs to offer us assurances that they are implementing. (GK)*
- *Including the letters of assurance that I already referenced, I think we'll know in large part about compliance, as there's this kind of word of mouth aspect as well. I mentioned our collaboration, our intentional collaboration with a variety of stakeholders, that they will let us know and already do let us know if a district hasn't reached joint agreement yet and hear their points of contention. What can we do to support them? I think there's also a potential that a community member or a teacher or a school leader or somebody in a district might pick up the phone and say, "Hey, we don't have a system in place that's meeting the intent of the law." In which case, the approach we would take is more of, "Let's take a look at your system and how we can help you improve it to come into compliance." (TL)*

### Lessons Learned from the Pilot of the Teacher Assessment Model

Several lessons were learned from the Pilot of the teacher assessment model. As previously mentioned, full and complete implementation will take a long time. In fact, it is work that will continually evolve. Models will be modified, incorporating technology in new ways and shifting to meet the ever-changing needs and goals of the educational system. It is expected that evaluation models will be tailored to specific schools and will be adjusted as roles and responsibilities become better defined.

To support the process of Model refinement, a focus on the big picture and deep understanding of the purpose should be emphasized in training and implementation. It is also important to continue to offer training and professional development to ensure continued fidelity of implementation and a strong emphasis on the overarching purposes of the Model.

A district's evaluation model should be integrated with overall goals and existing initiatives in order for the district to benefit fully from the time and effort put forth by school administration

and staff. A meaningful balance across all activities must be reached to prevent the participants from becoming overwhelmed.

It was also learned that peer observations may be a very difficult aspect of the Model for teachers. However, both MDE interviewees emphasized that peer observations have the potential to encourage collaboration and interdependence among teachers. It should be emphasized that even the best teachers can get better, and a community of trust must be developed before even the most well-intentioned constructive feedback can be taken and turned into professional growth rather than taken as a threat or insult.

In the Pilot year, the Model was applied to all persons on a teacher contract. Persons on teacher contracts included traditional classroom teachers as well as professionals whose work mostly takes place outside of a classroom (e.g., nurses, counselors, social workers). Feedback received from Pilot participants indicated that the Model did not provide a good fit for many non-classroom professionals. A different type of model may be more useful for these individuals, and districts might move toward developing a series of different evaluation models for different roles.

Overall, the Model requires a great deal of communication within and across groups and roles. Establishing new routines and expanded communication channels will take time, practice, and a belief that these will lead to increases in student performance. Changes in teacher practice will necessitate continuing discussions among teachers and principals to ensure that the changes result in improved education for all students.

### *Comments Regarding Lessons Learned from the Pilot of the Teacher Assessment Model*

- *I think teachers are buying into this work and that's why our partnership with the school board association and especially Education MN has been so important because they, and in most instances, teachers see this as a good thing and see this as their right. And, so I think when a district has been resistant, the teachers have been the ones who have said "we've got to do this" because they want that feedback. They want that development opportunity. (GK).*
- [How to assess counselors, nurse, school psychologists?] *I think in the end, most districts will move towards a series of different evaluation models for all of these different roles. I think they will have to do that. (GK)*
- *It* [the evaluation system] *will address the culture to get to the big ideas, rather than the "Did I do right?" aspect of the forms. (TL)*
- *The work is both hard and right. Teaching is a profession. It's a craft that's both technical and artistic, and having a system in place to measure everything between technical and artistic artisanship, is challenging, especially when time is always a consideration. It's worth the time only if and when it moves practice forward. The challenge, I feel like everyone's wrestling with, is at what threshold will we realize the results of moving practice forward? What will be the threshold of positive results for time invested in training, of learning, of supporting the school leader or the peers to offer feedback, and the teachers looking at the student data? Where is that magic spot where practice will move forward and*

*student outcomes will move forward after we've ascended this hill of getting through the forms and all steps that are necessary in this? (TL)*

## Summary of Findings of the MDE Staff Interviews

Individual interviews with two MDE staff members responsible for the development and implementation of the Pilot were conducted in July and August 2014. The staff members were Tyler Livingston, Educator Effectiveness Supervisor; and Greg, Keith, Director of the Division of School Support. The responses of the two interviewees are combined in the summary presented below.

### Successful Aspects of the Pilot Model Implementation

The interviewees described success as occurring on two levels. First, the overall Model appears to have successfully encouraged teacher buy-in to an evaluation process. Second, the Pilot presented the opportunity to learn which aspects of the Model were most effectively implemented and why.

The willingness of districts to participate in one or more components of the Model and their cooperation in carrying out required activities was cited as being critical to the success of the Pilot. The collaborative nature of the relationship between MDE and the Pilot districts was considered key.

The process of teachers and summative evaluators working together for a common purpose was perceived to have a very positive impact on teacher buy-in. Points of contact and feedback to teachers were considered to be the most successful individual elements of the Model.

### Areas for Future Development of the Model

The implementation of the Pilot was a learning experience that helped identify aspects of the Model where modifications would be beneficial. An idea that surfaced multiple times was having a phased implementation strategy, rather than rolling out all pieces of the Model at once. Both summative evaluators and teachers alike indicated they would have appreciated an initial introductory overview of the entire Model followed by step-by-step training for each individual component. A phased strategy such as this might have reduced confusion and feelings of being overwhelmed by the Model's expectations.

Possible revisions to specific aspects of the Model were also identified. For example, it was mentioned that the number of points of contact might be reduced. It was also suggested that the peer review process be analyzed to identify the training and ongoing coaching that teachers need in order to become effective as peer reviewers. The peer review process was considered to be one of the most valuable elements of the Model, and improved peer reviewer training could make it even more so. One surprising outcome was the difficulty that teachers had with the creation and measurement of student learning goals. It was suggested that increasing training on how to set, assess, and reflect on data from student learning goals would likely eliminate much of the difficulty.

A common theme in summative evaluator feedback was that implementation of the Model's multiple elements was a significant time commitment. When asked about this, both MDE interviewees noted that processes generally take longer when an initiative is first implemented than later on after it has become more routine. After the Model becomes familiar and summative evaluators and teachers have acquired experience with it, it is likely that the components will take less time and effort. It is also likely that summative evaluators will identify ways that the time commitment can be reduced without affecting the quality of the evaluation process

Several lessons were learned from the Pilot of the teacher assessment model. It is expected that evaluation models will continually evolve, incorporating technology in new ways and shifting to meet the ever-changing needs and goals of the educational system. It is expected that evaluation models will be tailored to specific schools and will be adjusted as roles and responsibilities become better defined.

To support the process of Model refinement, a focus on the big picture and deep understanding of the purpose should be emphasized in training and implementation. It is also important to continue to offer training and professional development to ensure continued fidelity of implementation and a strong emphasis on the overarching purposes of the Model.

A district's evaluation model should be integrated with overall goals and existing initiatives in order for the district to benefit fully from the time and effort put forth by school administration and staff. A meaningful balance across all activities must be reached to prevent the participants from becoming overwhelmed.

Overall, the Model requires a great deal of communication within and across groups and roles. Establishing new routines and expanded communication channels will take time, practice, and a belief that these will lead to increases in student performance. Changes in teacher practice will necessitate continuing discussions among teachers and principals to ensure that the changes result in improved education for all students.

# LIST OF APPENDICES

**Page**

| A | Caledonia (Whole Model) |
|---|---|
| B | Deer River (Student Engagement) |
| C | Detroit Lakes (Student Engagement) |
| D | Granada Huntley (Whole Model) |
| E | Lester Prairie (refer to ● on map) (Teacher Practice) |
| F | Montevideo (Student Engagement) |
| G | New Ulm (Whole Model) |
| H | North Shore Community School (Duluth) (Whole Model) |

| I | Orono (refer to the ● on the map) (Student Engagement) |
|---|---|
| J | Perpich Arts High School (refer to ● on the map) (Whole Model) |
| K | Pine Island (Whole Model) |
| L | Pine River Backus (Student Engagement & Student Learning Goals) |
| M | Prior Lake/Savage (Student Learning Goals) |
| N | Red Wing (Student Engagement & Student Learning Goals) |
| O | Saint James (Student Engagement) |
| P | Saint Peter (Student Engagement) |
| Q | Wabasha-Kellogg (Student Engagement & Student Learning Goals) |

Table B-1.  Descriptive information about the 17 districts (or schools) participating in the Pilot

| District or school | Total No. of schools | | | Total student enrollment | Total no. of teachers | Total no. of administrators |
|---|---|---|---|---|---|---|
| | Elementary | Middle | High school/Secondary | | | |
| Caledonia Public School District | 1 | - | 1 | 667 | 44.9 | 2.5 |
| Deer River Public School District | 1 | - | 1 | 897 | 66.5 | 3.3 |
| Detroit Lakes Public School District | 2 | 1 | 2[b] | 2,916 | 201.9 | 9 |
| Granada Huntley East Chain Public School District | 1 | - | 1 | 202 | 17.6 | 1.5 |
| Lester Prairie Public School District[a] | 1 | - | 1 | 418 | 29.5 | 2 |
| Montevideo Public School District | 2 | 1 | 1 | 1,321 | 87.9 | 3.5 |
| New Ulm Public School District | 2 | - | 1 | 1,942 | 127.6 | 5 |
| North Shore Community School | 1 | - | - | 333 | 20.9 | 1 |
| Orono Public School District | 2 | 1 | 1 | 2,794 | 147.7 | 10 |
| Perpich Center for Arts Education | - | - | 1 | 249 | 22.8 | 2 |
| Pine Island Public School District | 1 | 1 | 1 | 1,194 | 81.7 | 4.3 |
| Pine River Backus Public School District | 1 | - | 2[b] | 926 | 69.6 | 3.6 |
| Prior Lake/Savage Public School District | 7 | 2 | 2[b] | 7,510 | 391.2 | 21 |
| Red Wing Public School District | 2 | 1 | 1 | 2,746 | 171.1 | 9 |
| Saint James Public School District | 1 | - | 1 | 1,008 | 63.2 | 6 |
| Saint Peter Public School District | 2 | 1 | 1 | 1,955 | 124 | 6 |
| Wabasha-Kellogg Public School District | 1 | - | 1 | 588 | 43.6 | 2.8 |

[a]Grades K-6 and 7-12 are in the same building.
[b]Count includes an alternative learning center.
Note:  These 2013-14 data were obtained from Minnesota Department of Education website, http://rc.education.state.mn.us/. The table presents totals. In some districts, however, not all teachers and administrators were actively participating in the pilot.

## General Instructions

This survey is to be completed by teachers, or professional staff on a teacher's contract, in school districts that are piloting the Minnesota State Teacher Development, Evaluation, and Peer Support Model. **If you are not a classroom teacher, or if you are not on a teacher's contract (e.g., counselor, nurse), you do not need to complete this survey.**

Because school districts are piloting different components of the Model, you will be asked to complete only the parts of the survey that apply to your district. Here is an overview. (Note that this information will be repeated in the survey.)

**A. Teacher Practice - Individual Growth and Development Plan:** Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island

**B. Student Engagement:** Caledonia, Deer River, Detroit Lakes, Granada Huntley, Montevideo, North Shore, Orono, Perpich, Pine Island, Pine River-Backus, Red Wing, Saint James, Saint Peter, Wabasha-Kellogg

**C. Student Learning Goals:** Caledonia, Granada Huntley, New Ulm, North Shore, Perpich, Pine Island, Pine River-Backus, Prior Lake/Savage, Red Wing, Wabasha-Kellogg

**D. Overall Evaluation:** All Districts

**E. Background Information:** All Districts

**Please note that this survey might not work properly on smart phones or tablets.**

**✱1. Please indicate the type of position you hold. Refer to the explanations shown below.**

<u>Classroom teachers</u> instruct groups of students such as kindergarten, grade 1, high school biology, or middle school music. Classroom teachers include reading specialists, EL/ELL, special education.

<u>Non-teaching professionals on a teacher contract</u> include counselors, school nurses, social workers, librarians, and other non-teaching professional positions.

If your work falls into a different category, please select <u>Other</u>.

○ I am a *classroom* teacher.

○ I am *non-teaching professional staff on a teacher contract*.

○ Other, I am *not* a teacher and I am *not on a teacher contract*.

## A. Teacher Practice - Individual Growth and Development Plan

Part A to be completed by: **Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island**

**2. If you are <u>not</u> one of these districts, click Go to Part B and Next. Otherwise, just click Next.**

○ Go to Part B

## A (continued) Teacher Practice - Individual Growth and Development Plan

Part A questions are related to your Individual Growth and Development Plan (the "Plan"). We would like your assistance in evaluating the elements included in the Plan.

### 3. Did you complete a Plan for 2013-2014?

○ Yes

○ No

### 4. How much time did you spend preparing your Plan? (For example, if you spent 1 hour and 20 minutes to complete your plan, enter Hours: 1 and Minutes: 20.)

Hours [          ]

Minutes [          ]

### 5. When working on your Plan, did you refer to these items?

|  | Yes | No |
|---|---|---|
| District goals | ○ | ○ |
| School goals | ○ | ○ |
| Professional learning community (PLC) goals | ○ | ○ |

### 6. In formulating my Plan, I consulted with

|  | Yes | No |
|---|---|---|
| A peer reviewer | ○ | ○ |
| A summative evaluator | ○ | ○ |
| Other | ○ | ○ |
| No one; I wrote them on my own | ○ | ○ |

If other, please specify

[                                        ]

### 7. In formulating my Plan, it was beneficial to consult with:

|  | Yes | No | N/A |
|---|---|---|---|
| A peer reviewer | ○ | ○ | ○ |
| A summative evaluator | ○ | ○ | ○ |
| Other | ○ | ○ | ○ |

### 8. The next two items concern writing your professional growth goals and performance standards.

|  | Very difficult | Somewhat difficult | Neither easy nor difficult | Somewhat easy | Very easy |
|---|---|---|---|---|---|
| I found writing my professional growth goal(s) to be... | ○ | ○ | ○ | ○ | ○ |
| I found writing the performance standard(s) for my goal(s) to be... | ○ | ○ | ○ | ○ | ○ |

**9. Please answer the following regarding your Plan.**

| | Completely untrue | Somewhat untrue | Neither untrue nor true | Somewhat true | Completely true | N/A |
|---|---|---|---|---|---|---|
| The *Handbook* provided adequate guidance for developing my Plan. | ○ | ○ | ○ | ○ | ○ | ○ |
| MDE training provided adequate guidance for developing my Plan. | ○ | ○ | ○ | ○ | ○ | ○ |
| Formulating professional growth goals helped me to reflect on my instructional practice. | ○ | ○ | ○ | ○ | ○ | ○ |

**10. How would you rate the Individual Growth and Development Plan with respect to generating information that enables your summative evaluator to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to your performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where you could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## A (continued). Teacher Practice - Points of Contact

**11. How many points of contact between you and <u>peer reviewer(s)</u> and a <u>summative evaluator</u> occurred during the 2013-2014 school year?**

Points of contact with <u>peer reviewer(s)</u>

Points of contact with a <u>summative evaluator</u>

**12. How would you rate Points of Contact with respect to generating information that enables your summative evaluator to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to your performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where you could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## A (continued). Performance Standards for Teacher Practice

**13. The *Performance Standards for Teacher Practice Rubric* was an effective tool for...**

| | Completely untrue | Somewhat untrue | Neither untrue nor true | Somewhat true | Completely true | N/A |
|---|---|---|---|---|---|---|
| Identifying areas of my practice for professional growth. | ○ | ○ | ○ | ○ | ○ | ○ |
| Identifying areas of my practice for evaluation. | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## B. Student Engagement

Part B to be completed by: **Caledonia, Deer River, Detroit Lakes, Granada Huntley, Montevideo, North Shore, Orono, Perpich, Pine Island, Pine River-Backus, Red Wing, Saint James, Saint Peter, Wabasha-Kellogg**

**14. If you are <u>not</u> one of these districts, click Go to Part C and Next. Otherwise, just click Next.**

○ Go to Part C

## B (continued) Student Engagement - YouthTruth Student Survey

**15. Did <u>your</u> students complete the YouthTruth survey this year?**

○ Yes

○ No

**16. What methods (other than the YouthTruth survey) did you use to assess student engagement?**

☐ Student survey other than the YouthTruth survey

☐ Observation

☐ Student portfolios

☐ Student presentations/performances/projects

☐ Teacher portfolio

☐ Video

☐ Other

If other, please specify

**17. How much time did you spend working on the assessment of student engagement? (For example, if you spent a total of 1 hour and 20 minutes, enter Hours: 1 and Minutes: 20.)**

Hours

Minutes

**18. How would you rate Student Engagement assessments with respect to generating information that enables your summative evaluator to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to your performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where you could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## C. Student Learning Goals

Part C to be completed by: **Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island, Pine River-Backus, Prior Lake/Savage, Red Wing, Wabasha-Kellogg**

**19. If you are _not_ one of these districts, click Go to Part D and Next. Otherwise, just click Next.**

○ Go to Part D

## C (continued). Student Learning Goals

**20. Please answer the following regarding Student Learning Goals.**

| | Completely untrue | Somewhat untrue | Neither untrue nor true | Somewhat true | Completely true | N/A |
|---|---|---|---|---|---|---|
| The Handbook provided adequate guidance for developing my student learning goals. | ○ | ○ | ○ | ○ | ○ | ○ |
| The MDE training provided adequate guidance for developing my student learning goals. | ○ | ○ | ○ | ○ | ○ | ○ |

**21. The following items concern assessments, mastery scores, starting points, and setting student learning goals.**

| | Very difficult | Somewhat difficult | Neither easy nor difficult | Somewhat easy | Very easy | N/A |
|---|---|---|---|---|---|---|
| I found selecting or developing an **assessment** to measure the goals to be... | ○ | ○ | ○ | ○ | ○ | ○ |
| I found establishing **mastery scores** for the assessment to be... | ○ | ○ | ○ | ○ | ○ | ○ |
| I found establishing student **starting points** to be... | ○ | ○ | ○ | ○ | ○ | ○ |
| I found setting student **learning goals** to be... | ○ | ○ | ○ | ○ | ○ | ○ |

**22. In formulating my student learning goals, I consulted with:**

☐ A peer reviewer

☐ A summative evaluator

☐ Other

☐ No one; I wrote them on my own

If other, please specify

[ ]

**23. In formulating my student learning goals, it was beneficial to consult with:**

| | Yes | No | N/A |
|---|---|---|---|
| A peer reviewer | ○ | ○ | ○ |
| A summative evaluator | ○ | ○ | ○ |
| Other | ○ | ○ | ○ |

If other, please specify

[ ]

**24. How much time did you spend working on your student learning goals? (For example, if you spent 1 hour and 20 minutes, enter Hours: 1 and Minutes: 20.)**

Hours [ ]

Minutes [ ]

**25. How would you rate Student Learning Goals with respect to generating information that enables your summative evaluator to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to your performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where you could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

## D. Overall Evaluation

**All Districts**

**26. Based on your experiences this year, how how would you rate...**

|  | Poor | Fair | Good | Very good | Excellent | No opinion |
|---|---|---|---|---|---|---|
| the Minnesota State Teacher Development, Evaluation, and Peer Support Model overall? | ○ | ○ | ○ | ○ | ○ | ○ |
| the Model's ability to provide a valid assessment of your performance? | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

**27. Now that you have acquired more experience working with the Model, has your opinion about the Model changed?**

○ Yes

○ No

If **YES**, please explain:

**28. How might the Model be changed so that it would be more effective? (Check all that apply)**

☐ **Simplify**; It's too complex.

☐ **Clarify**; It's difficult to follow the instructions or see how the parts fit together.

☐ **Streamline**; It's too time consuming.

☐ **Change the component weights**; current weights are .45 Teacher Practice, .20 Student Engagement, .35 Student Learning and Achievement.

☐ **Change the number of rating categories**; current categories are Exemplary, Effective, Development Needed, Unsatisfactory.

☐ **Other**

Comments:

**29. Based on your experience, what would you say are the <u>benefits</u> associated with the use of the Model? (Check all that apply)**

- ☐ Incorporates multiple sources of information
- ☐ Facilitates evaluator-to-teacher conversations regarding instructional practice
- ☐ Facilitates teacher-to-teacher conversations regarding instructional practice
- ☐ Provides information on how teachers can improve their instructional practice
- ☐ Enables teachers to see improvement over time
- ☐ Informs district decisions regarding professional development opportunities
- ☐ Other

Comments:

**30. Please use the space below to write any comments or suggestions you have regarding the Minnesota State Teacher Development, Evaluation, and Peer Support Model.**

## E. Background Informaton

**All Districts**

**31. Please select your district:**

**32. Grades you are teaching this school year in your district (check all that apply)**

| | | |
|---|---|---|
| ☐ PreK | ☐ 4 | ☐ 9 |
| ☐ K | ☐ 5 | ☐ 10 |
| ☐ 1 | ☐ 6 | ☐ 11 |
| ☐ 2 | ☐ 7 | ☐ 12 |
| ☐ 3 | ☐ 8 | ☐ Other |

**33. Your teaching role. (Please select those that apply.)**

☐ PreK

☐ Elementary Generalist

☐ Elementary Specialist (e.g., Arts, PE, World Languages)

☐ Secondary Core (Language Arts, Sciences, Math, Social Studies)

☐ Secondary Elective (e.g., Ag, Business, World Languages)

☐ EL/ELL

☐ Special Education

Other (please specify)

[                                              ]

**34. Subjects you taught this year**

[                                              ]

**35. Your total number of years teaching (do not include this year)**

[          ▼]

If more than 20, please specify

[          ]

**Thank you for participating in this survey.**

Table B-1. Classroom and non-classroom teachers' total number of years taught.

| Descriptive statistics for number of years taught | Classroom teachers (*n=270*) | Non-classroom teachers (*n=12*) |
|---|---|---|
| Mean number of years taught | 16.5 | 14.3 |
| Standard deviation of years taught | 10.2 | 10.8 |
| Median number of years taught | 16 | 12.5 |
| Range of years taught | 0-44 | 0-34 |

Table B-2. Number of classroom and non-classroom teachers at each grade level.

| Grade level | Classroom teachers (*n=273*) | Non-classroom teachers (*n=11*) |
|---|---|---|
| **Pre-K** | **12** | **2** |
| **K-5** | **119** | **7** |
| Kindergarten | 41 | 5 |
| Grade 1 | 34 | 5 |
| Grade 2 | 37 | 4 |
| Grade 3 | 41 | 3 |
| Grade 4 | 34 | 4 |
| Grade 5 | 46 | 4 |
| **Middle school (6-8)** | **105** | **3** |
| Grade 6 | 46 | 3 |
| Grade 7 | 54 | 2 |
| Grade 8 | 61 | 2 |
| **High school (9-12)** | **124** | **2** |
| Grade 9 | 75 | 2 |
| Grade 10 | 77 | 2 |
| Grade 11 | 105 | 2 |
| Grade 12 | 101 | 2 |
| **Other** | **5** | **3** |

Table B-3. Number of classroom teacher and non-classroom teachers by role.

| Role | Classroom teachers (*n=275*) | Non-classroom teachers (*n=11*) |
|---|---|---|
| Pre-K | 8 | 0 |
| Elementary generalist | 82 | 0 |
| Elementary specialist | 25 | 3 |
| Secondary core | 82 | 0 |
| Secondary elective | 53 | 0 |
| EL/ELL | 4 | 0 |
| Special education | 39 | 2 |
| Other | 14 | 9 |

## Welcome to the Summative Evaluator Survey

This survey is to be completed by summative evaluators in school districts that are piloting the Minnesota State Teacher Development, Evaluation, and Peer Support Model. **If you are not a summative evaluator in your district, please do not complete the survey.**

Because school districts are piloting different components of the Model, you will be asked to complete only the parts of the survey that apply to your district. Here is an overview. (Note that this information will be repeated in the survey.)

**A. Information about Your Building:** All Districts
**B. Teacher Practice - Individual Growth and Development Plans:** Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island
**C. Teacher Practice - Points of Contact:** Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island
**D. Student Learning Goals:** Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island, Pine River-Backus, Prior Lake/Savage, Red Wing, Wabasha-Kellogg
**D. YouthTruth Student Engagement Survey:** Caledonia, Deer River, Detroit Lakes, Granada Huntley, Montevideo, New Ulm, North Shore, Orono, Perpich, Pine Island, Pine River Backus, Red Wing, Saint James, Saint Peter, Wabasha-Kellogg
**F. MDE Communication and Training:** All Districts
**G. Satisfaction with the Model, Communication, and Training:** All Districts

**Please note that this survey might not work properly on all smart phones and tablets.**

### 1. Did you serve as a summative evaluator this year (2013-14)?

◯ Yes

◯ No

## A. Information about Your Building

Part A to be completed by: **All Districts**

**If you are serving as a summative evaluator in only one building in your district, please answer Questions 2-5 for just that building. If you are serving as a summative evaluator in more than one building, please answer Questions 2-5 in terms of ALL those buildings combined.**

### 2. How many summative evaluators are there in the building(s) for which you are serving as a summative evaluator?

Enter the total number of summative evaluators.

### 3. How many teachers in those building(s) are participating in the Pilot this year (2013-2014)?

Enter the total number of teachers.

### 4. For how many teachers are you serving as the summative evaluator this year (2013-14)?

Enter the number of teachers.

**5. What are the grade level assignments of the teachers for whom you are serving as summative evaluator? (Check all that apply)**

| | | |
|---|---|---|
| ☐ PreK | ☐ 4 | ☐ 9 |
| ☐ K | ☐ 5 | ☐ 10 |
| ☐ 1 | ☐ 6 | ☐ 11 |
| ☐ 2 | ☐ 7 | ☐ 12 |
| ☐ 3 | ☐ 8 | ☐ Other |

If other, please specify

## B. Teacher Practice - Individual Growth and Development Plans

Part B to be completed by: **Caledonia, Granada Huntley, Lester Prairie, New Ulm, North Shore, Perpich, Pine Island**

**6. If you are _not_ in one of these districts, click 'Go to Part D' and 'Next'. Otherwise, just click 'Next'.**

○ Go to Part D

## B. Teacher Practice - Individual Growth and Development Plans (continued)

**7. How many teachers' Individual Growth and Development Plans did you approve this year?**

Enter the number of teachers.

**8. How much time did you spend on average, PER TEACHER, reviewing and approving Individual Growth and Development Plans this year? (For example, if you spent, on average, 1 hour and 10 minutes per teacher, enter Minutes: 70)**

Minutes

**9. How would you rate the teachers' Individual Growth and Development Plans with respect to generating information that enables you to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where a teacher could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## C. Teacher Practice - Points of Contact

**10. For how many teachers did you carry out points of contact this year?**

Enter the number of teachers. [_____]

**11. How much time did you spend on average, PER TEACHER, preparing for pre-observation points of contact? (For example, if you spent, on average, 1 hour and 10 minutes per teacher, enter Minutes: 70)**

Minutes [_____]

**12. How much time did you spend on average, PER TEACHER, in the actual observation points of contact (e.g., classroom observation)?**

Minutes [_____]

**13. How much time did you spend on average PER TEACHER, in activities related to post-observation points of contact?**

Minutes [_____]

**14. How would you rate the points of contact with respect to generating information that enables you to...**

|  | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where a teacher could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

[                                                    ]

## D. Student Learning Goals

Part D to be completed by: **Caledonia, Granada Huntley, New Ulm, North Shore, Perpich, Pine Island, Pine River-Backus, Prior Lake/Savage, Red Wing, Wabasha-Kellogg**

**15. If you are not in one of these districts, click 'Go to Part E' and 'Next'. Otherwise, just click 'Next'.**

○ Go to Part E

## D. Student Learning Goals (continued)

**16. How many teachers' Student Learning Goals forms did you approve this year?**

Number of teachers [_____]

**17. How much time did you spend on average, PER TEACHER, working on Student Learning Goals? (For example, if you spent, on average, 1 hour and 10 minutes per teacher, enter Minutes: 70)**

Minutes [                    ]

**18. How would you rate Student Learning Goals with respect to generating information that enables you to...**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where a teacher could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:
[                                                                    ]

## E. YouthTruth Student Engagement Survey

Part E to be completed by: **Caledonia, Deer River, Detroit Lakes, Granada Huntley, Montevideo, New Ulm, North Shore, Orono, Perpich, Pine Island, Pine River Backus, Red Wing, Saint James, Saint Peter, Wabasha-Kellogg.**

**19. Did you participate in training provided by YouthTruth regarding student surveys?**

○ Yes

○ No

**20. Overall, how satisfied are you with...**

| | Very dissatisfied | Dissatisfied | Neither dissatisfied nor satisfied | Satisfied | Very satisfied | N/A |
|---|---|---|---|---|---|---|
| the training you received from YouthTruth? | ○ | ○ | ○ | ○ | ○ | ○ |
| the support, other than training, you received from YouthTruth? | ○ | ○ | ○ | ○ | ○ | ○ |
| the process of administering the YouthTruth survey to students in your district? | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:
[                                                                    ]

**21. How many of the <u>teachers</u> for whom you served as summative evaluator had their students take the YouthTruth survey?**

Number of teachers [                    ]

**22. Other than training sessions, what was is the <u>TOTAL AMOUNT OF TIME</u> you spent on activities related to the YouthTruth survey? (For example, if you spent a TOTAL of 1 hour and 40 minutes, enter Minutes: <u>100</u>)**

Minutes [          ]

**23. How would you rate the YouthTruth student surveys with respect to generating information that enables you to:**

|  | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where a teacher could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:
[                    ]

## F. MDE Communication and Training

Part F to be completed by: **All Districts**

**24. Overall, how satisfied are you with...**

|  | Very dissatisfied | Dissatisfied | Neither dissatisfied nor satisfied | Satisfied | Very satisfied | N/A |
|---|---|---|---|---|---|---|
| the **communication** you have had with MDE staff regarding the Pilot? | ○ | ○ | ○ | ○ | ○ | ○ |
| the **training** you have received from MDE for the Pilot? | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:
[                    ]

## G. Satisfaction with the Model

Part H to be completed by: **All Districts**

### 25. From your perspective, how well does the Model seem to work for evaluating...

| | Poor | Fair | Good | Very good | Excellent | N/A |
|---|---|---|---|---|---|---|
| Probationary teachers? | ○ | ○ | ○ | ○ | ○ | ○ |
| Mid-career teachers? | ○ | ○ | ○ | ○ | ○ | ○ |
| Late-career teachers? | ○ | ○ | ○ | ○ | ○ | ○ |
| Specialist teachers? | ○ | ○ | ○ | ○ | ○ | ○ |
| Part-time teachers? | ○ | ○ | ○ | ○ | ○ | ○ |
| Teachers who teach in more than one school in the district? | ○ | ○ | ○ | ○ | ○ | ○ |

### 26. Based on your experiences this year, how how would you rate...

| | Poor | Fair | Good | Very good | Excellent | No opinion |
|---|---|---|---|---|---|---|
| the Minnesota State Teacher Development, Evaluation, and Peer Support Model overall? | ○ | ○ | ○ | ○ | ○ | ○ |
| the Model's ability to provide a valid assessment of a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

### 27. Now that you have acquired more experience working with the Model, has your opinion about the Model changed?

○ Yes

○ No

If **YES**, please explain:

**28. How might the Model be changed so that it would be more effective? (Check all that apply)**

☐ **Simplify**; It's too complex.

☐ **Clarify**; It's difficult to follow the instructions or see how the parts fit together.

☐ **Streamline**; It's too time consuming.

☐ **Change the component weights**; current weights are .45 Teacher Practice, .20 Student Engagement, .35 Student Learning and Achievement.

☐ **Change the number of rating categories**; current categories are Exemplary, Effective, Development Needed, Unsatisfactory.

☐ **Other**

Comments:

[ text box ]

**29. Based on your experience, what would you say are the <u>benefits</u> associated with the use of the Model? (Check all that apply)**

☐ Incorporates multiple sources of information

☐ Facilitates evaluator-to-teacher conversations regarding instructional practice

☐ Facilitates teacher-to-teacher conversations regarding instructional practice

☐ Provides information on how teachers can improve their instructional practice

☐ Enables teachers to see improvement over time

☐ Informs district decisions regarding professional development opportunities

☐ Other

Comments:

[ text box ]

**30. Please use the space below to write any comments or suggestions you have about the Minnesota State Teacher Development, Evaluation, and Peer Support Model.**

[ text box ]

**Thank you for taking time to complete our survey.**

## TEACHER BELIEFS ABOUT STUDENT ENGAGEMENT MEASURES SURVEY

The purpose of this survey is to gather information on teachers' beliefs about the value of student engagement in evaluating teacher practice. Teacher input is critical to the development of student engagement measures so that those measures can be authentic and valued by teachers. The survey results will be used by MDE to shape the teacher development, evaluation and peer support model that will eventually be made available to districts across the state.

Your individual survey responses will be kept confidential. Only researchers at the University of Minnesota's Center for Applied Research and Educational Improvement (CAREI) will have access to your individual survey results. CAREI will combine your responses with responses from other Minnesota teachers and provide an aggregate report to MDE.

The survey will be repeated next spring to determine how teachers' beliefs might have changed during their participation in the pilot. We ask you to provide your name so that we can match your responses on the two surveys.

If you have any questions about this survey or CAREI's study of the *Minnesota State Teacher Development, Evaluation, and Peer Support Model* pilot, feel free to contact Deb Ingram at d-ingram@umn.edu or 612.625.0502.

## TEACHER BELIEFS ABOUT STUDENT ENGAGEMENT MEASURES SURVEY

As part of the pilot for the *Minnesota State Teacher, Development, Evaluation, and Peer Support Model,* schools will be administering a student survey to gather student feedback about their experiences with specific teachers related to seven constructs: academic engagement (extent to which students are meeting grade level expectations), behavioral engagement (level students are participating in class and are on-task), affective engagement (students feel connected and identify with the class), cognitive engagement (self-regulating and see value in learning), relationship building, rigorous instruction, and relevant instruction. Teachers and evaluators will get the results of the student survey for use in teacher development and evaluation.

NAME (First, Last) _____

SCHOOL NAME(S) _____

SCHOOL DISTRICT _____

### A. USEFULNESS OF A STUDENT SURVEY TO MEASURE ENGAGEMENT

**DIRECTIONS.** Please indicate your level of agreement or disagreement with each of the statements below by filling in the appropriate oval.

| | STRONGLY DISAGREE | DISAGREE | SOMEWHAT DISAGREE | NEITHER DISAGREE NOR AGREE | SOMEWHAT AGREE | AGREE | STRONGLY AGREE |
|---|---|---|---|---|---|---|---|
| 1. Student surveys will provide me with feedback that will enable me to identify strengths and weaknesses in my teaching methods. | O | O | O | O | O | O | O |
| 2. Student surveys will provide me with information that will enable me to improve my effectiveness as a teacher. | O | O | O | O | O | O | O |
| 3. Responses that my students give on the student survey can accurately predict how well they do on learning performance measures. | O | O | O | O | O | O | O |
| 4. I am in favor of using feedback from student surveys to evaluate my teaching. | O | O | O | O | O | O | O |
| 5. Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors. | O | O | O | O | O | O | O |
| 6. Student surveys are not worth administering because I can easily influence students so they give positive ratings. | O | O | O | O | O | O | O |
| 7. It will take a lot of time for me to review the results of my students' surveys in order to make them useful. | O | O | O | O | O | O | O |
| 8. Feedback from student surveys can help my school identify areas where teachers need professional development. | O | O | O | O | O | O | O |
| 9. Feedback from student surveys can be used to track changes in my teaching performance over time. | O | O | O | O | O | O | O |
| 10. Surveying my students is worth the time it takes to administer the survey. | O | O | O | O | O | O | O |
| 11. I am concerned that students will give low ratings in a subject that is difficult for them. | O | O | O | O | O | O | O |
| 12. I am concerned that students will use the survey to get back at teachers. | O | O | O | O | O | O | O |

**Continue on other side** ⇨

## TEACHER BELIEFS ABOUT STUDENT ENGAGEMENT MEASURES SURVEY

### B. INFORMATION ABOUT TEACHER RESPONDENT

13. Grade(s) you will teach this school year.
    (Fill in all that apply):

    O    K
    O    1
    O    2
    O    3
    O    4
    O    5
    O    6
    O    7
    O    8
    O    9
    O    10
    O    11
    O    12
    O    Other: _____

14. Please indicate your total number of years teaching *in this district* (Do *not* include this year):

    ⓪   ⓪
    ①   ①
    ②   ②
    ③   ③
    ④   ④
    ⑤   ⑤
        ⑥
        ⑦
        ⑧
        ⑨

15. Subject(s) you will be teaching this school year.

    _____

    _____

    _____

16. Please indicate your total number of years teaching (Do *not* include this year):

    ⓪   ⓪
    ①   ①
    ②   ②
    ③   ③
    ④   ④
    ⑤   ⑤
        ⑥
        ⑦
        ⑧
        ⑨

17. I am:

    O    Female
    O    Male
    O    Other: _____

## Survey Description

As part of the Student Engagement component of the Minnesota State Teacher Development, Evaluation, and Peer Support Pilot, schools administer a student survey to gather feedback from students about their experiences with specific teachers. Teachers and evaluators then receive a summary of the student survey results for use in teacher development and evaluation. The purpose of the survey you are taking right now is to gather information about teachers' perceptions regarding the use of student surveys.

**1. Are you a classroom teacher? (i.e., Do you instruct a group of students? Classroom teachers include reading specialists, EL/ELL, special education, but *does not include* counselors, school nurses, and the like).**

○ Yes, I am a classroom teacher.

○ No, I am not a classroom teacher. (Completing this survey is optional)

**2. What is your name?**

FIRST _____

LAST _____

**3. What is your school district?**

○ Caledonia        ○ New Ulm                      ○ Prior Lake/Savage

○ Deer River       ○ North Shore Community Sch    ○ Red Wing

○ Detroit Lakes    ○ Orono                        ○ Saint James

○ Granada-Huntley  ○ Perpich Arts High School     ○ Saint Peter

○ Lester Prairie   ○ Pine Island                  ○ Wabasha-Kellogg

○ Montevideo       ○ Pine River-Backus

**4. What school(s) did you teach in this year?**

_____

## Usefulness of the Student Survey for Measuring Engagement

## 5. DIRECTIONS. Please indicate your level of agreement or disagreement with each of the statements below by filling in the appropriate oval.

| | Strongly disagree | Disagree | Somewhat disagree | Neither disagree nor agree | Somewhat agree | Agree | Strongly agree |
|---|---|---|---|---|---|---|---|
| a. Student surveys will provide me with feedback that will enable me to identify strengths and weaknesses in my teaching methods. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| b. Student surveys will provide me with information that will enable me to improve my effectiveness as a teacher. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| c. Responses that my students give on the student survey can accurately predict how well they do on learning performance measures. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| d. I am in favor of using feedback from student surveys to evaluate my teaching. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| e. Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| f. Student surveys are not worth administering because I can easily influence students so they give positive ratings. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| g. It will take a lot of time for me to review the results of my students' surveys in order to make them useful. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| h. Feedback from student surveys can help my school identify areas where teachers need professional development. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| i. Feedback from student surveys can be used to track changes in my teaching performance over time. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| j. Surveying my students is worth the time it takes to administer the survey. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| k. I am concerned that students will give low ratings in a subject that is difficult for them. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| l. I am concerned that students will use the survey to get back at teachers. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

## 6. Since the beginning of the school year, my opinion about the use of student surveys for teacher assessment has...

○ Remained about the same.

○ Become less favorable.

○ Become more favorable.

If you opinion *has* changed, why?

[ text entry box ]

## YouthTruth Survey Implementation

**7. Did _your_ students complete the YouthTruth survey this year?**

○ Yes

○ No

**8. Have you seen the summary report of your students' YouthTruth survey results?**

○ Yes

○ No

**9. Have you discussed the YouthTruth survey results with a summative evaluator yet?**

○ Yes

○ No

**10. How would you rate the YouthTruth student surveys with respect to generating information that enables a school to . . .**

| | Poor | Fair | Good | Very Good | Excellent | No Opinion | N/A |
|---|---|---|---|---|---|---|---|
| assign a valid rating to a teacher's performance? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Identify areas where a teacher could benefit from professional development? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Comments:

## Demographic Data

This information will help us understand perceptions of teachers with respect to grades and subjects taught. Individual respondents will never be identified.

**11. Grade(s) you taught this school year (Select all that apply):**

☐ PreK        ☐ 4         ☐ 9

☐ K          ☐ 5         ☐ 10

☐ 1          ☐ 6         ☐ 11

☐ 2          ☐ 7         ☐ 12

☐ 3          ☐ 8

Other teaching assignments in this district:

**12. Your teaching role. (Please select those that apply.)**

- [ ] PreK
- [ ] Elementary Generalist
- [ ] Elementary Specialist (e.g., Arts, PE, World Languages)
- [ ] Secondary Core (Language Arts, Sciences, Math, Social Studies)
- [ ] Secondary Elective (e.g., Ag, Business, World Languages)
- [ ] EL/ELL
- [ ] Special Education
- [ ] Other (please specify)

**13. Subject(s) you taught this school year.**

Center for Applied Research and Educational Improvement, College of Education and Human Development
University of Minnesota

Teacher Beliefs about Student Engagement Measures Survey

Table F-1.  Fall to spring change in teachers' ratings of statements that express support for the use of student surveys for teacher evaluation

| Survey item | n | Survey | % selecting response option on fall and spring surveys | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Strongly disagree | Disagree | Somewhat disagree | Neither disagree nor agree | Somewhat agree | Agree | Strongly Agree |
| I am in favor of using feedback from student surveys to evaluate my teaching | 218 | Fall | 5% | 8% | 16% | 16% | 26% | 23% | 6% |
| | | Spring | 12% | 12% | 17% | 14% | 26% | 15% | 4% |
| Student surveys will provide me with feedback that will enable me to identify strengths and weaknesses in my teaching methods | 219 | Fall | 3% | 4% | 9% | 7% | 33% | 37% | 7% |
| | | Spring | 7% | 10% | 7% | 10% | 38% | 26% | 2% |
| Student surveys will provide me with information that will enable me to improve my effectiveness as a teacher | 219 | Fall | 3% | 4% | 6% | 9% | 33% | 37% | 8% |
| | | Spring | 6% | 10% | 8% | 10% | 43% | 21% | 3% |
| Feedback from student surveys can be used to track changes in my teaching performance over time | 219 | Fall | 3% | 12% | 18% | 16% | 33% | 17% | 1% |
| | | Spring | 9% | 18% | 12% | 22% | 27% | 12% | 1% |
| Feedback from student surveys can help my school identify areas where teachers need professional development | 219 | Fall | 2% | 6% | 15% | 17% | 37% | 20% | 3% |
| | | Spring | 8% | 17% | 10% | 13% | 36% | 16% | 1% |
| Surveying my students is worth the time it takes to administer the survey | 217 | Fall | 3% | 12% | 10% | 20% | 29% | 22% | 4% |
| | | Spring | 12% | 11% | 12% | 16% | 26% | 20% | 3% |

Table F-2. Fall to spring change in teachers' ratings of statements that express concerns regarding the use of student surveys for teacher evaluation

| Survey item | *n* | Survey | % selecting response option on fall and spring surveys | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Strongly disagree | Disagree | Somewhat disagree | Neither disagree nor agree | Somewhat agree | Agree | Strongly Agree |
| I am concerned that students will give low ratings in a subject that is difficult for them | 217 | Fall | 2% | 7% | 8% | 14% | 33% | 27% | 9% |
| | | Spring | 1% | 6% | 7% | 12% | 35% | 24% | 16% |
| I am concerned that students will use the survey to get back at teachers | 217 | Fall | 2% | 9% | 9% | 20% | 35% | 17% | 9% |
| | | Spring | 1% | 7% | 7% | 18% | 35% | 19% | 13% |
| Student surveys are not worth administering because I can easily influence students so they give positive ratings | 219 | Fall | 6% | 22% | 18% | 25% | 22% | 5% | 3% |
| | | Spring | 6% | 16% | 15% | 31% | 19% | 10% | 4% |
| Students in the grades and subjects that I teach can discriminate between effective and ineffective teaching behaviors | 215 | Fall | 6% | 14% | 23% | 13% | 33% | 8% | 3% |
| | | Spring | 12% | 16% | 24% | 13% | 24% | 10% | 1% |
| Responses that my students give on the student survey can accurately predict how well they do on learning performance measures | 217 | Fall | 3% | 18% | 21% | 24% | 30% | 5% | 0% |
| | | Spring | 12% | 17% | 25% | 22% | 20% | 5% | 1% |
| It will take a lot of time for me to review the results of my students' surveys in order to make them useful | 217 | Fall | 3% | 20% | 15% | 24% | 25% | 12% | 1% |
| | | Spring | 6% | 23% | 23% | 24% | 17% | 7% | 1% |

# Interview Questions for Summative Evaluators – Spring/Summer 2014

1. Overall, how was this process for you?

*Teacher Practice*

2. How well do you think the following generated fair evidence for summative evaluations? Why?
   - Points of Contact
   - Portfolio
   - Teacher Practice Component Rating

3. How useful were the following in planning ongoing professional development? How so?
   - Points of Contact
   - Portfolio
   - Teacher Practice Component Rating

4. What challenges have you experienced in doing Points of Contact? Benefits?

5. Which of the following components in the Model should be maintained for future teacher development evaluations?
   - Performance Standards for Teacher Practice (rubric)?
   - Points of Contact?
   - Self-Assessment and Peer Review? Teacher Portfolio?

6. Are there any types of teachers for whom the Model is not working as well? If yes, which teachers and **what is not working well for them**? [Types of teachers include: specialist, probationary, mid-career, late-career, part-time, and teachers who teach in more than one school in the district.]

*Student Engagement (grades 3-12)*

1. How did the Student Engagement Component go this year? How did you work with your teachers to help them interpret and use the results?

2. How did you incorporate the Student Engagement survey results and other measures in developing the student engagement component rating?

*Student Learning Goals*

1. How successful do you think your teachers were at picking goals, selecting assessments, and interpreting results? Can you give me an example where a teacher was successful.  What did it look like?  An example of a teacher who struggled?

2. What did you do to determine consistent component ratings?

3. How did you turn assessment results into conversations about development and evaluation?

*Whole Model*

1. What effect, if any, did implementing the pilot have on your school culture?

2. How did implementing the Pilot affect students, if at all?

3. To what extent are your teachers changing their practice based their experience with the Model (planning, instruction, environment, professionalism)?

4. Which components of the Model most fairly differentiate effective teachers and teachers needing support?

5. Did you provide a final performance rating for any of your teachers this spring? (summative performance rating) If no, do you Plan to and when? If yes, how did it go? (surprises?) How useful was it to have an overall rating?

*Overall*

1. Given what you know at this point in the Pilot, are there ways you Plan to alter your approach to teacher development and evaluation next year? If so, how and in what ways?

2. Is there anything else you would like to tell us about your experience with the Model?

*[Thank them for completing the interview. Remind them to complete the online survey if they haven't already done it.*

***Minnesota State Teacher Development, Evaluation, and Peer Support Model* Pilot**
**Teacher Interview Questions – Spring data collection**

Thank you for agreeing to participate in this interview. The purpose of the interview is to collect information about how it has gone this year implementing the *Minnesota State Teacher Development, Evaluation, and Peer Support Model* ("the Model") This is very complex work so feedback from individual teachers like you is critical to helping MDE understand how the Model is working and how it can potentially be improved.

We appreciate your candor. CAREI researchers will combine the information from your interview with information from other interviews in this district in addition to other districts across the state.

We will not use your name or the name of your school or district in any written or verbal reports based on this data. We may include a direct quote from your interview, but we would not include any information that could be used to identify you, your school, or your district.

May I record this interview? No one but members of the CAREI research team will have access to this recording or the transcript of the recording. All recordings and transcripts will be stored in locked offices in our center.

Do you have any questions for me before I begin?

*Interviewer Instructions: Start the recorder and then state your name, the date, the school and the name of the interviewee before you ask the first question.*

*In case you need to show the interviewee what you're talking about, a blank Individual Growth and Development Plan, a blank Point of Contact Documentation form, a blank Student Learning Goals Documentation form, and the Performance Standards for Teacher Practice Rubric are attached at the end of this interview.*

*General*
   1. Overall, how was this process for you? (participating in the Teacher Development, Evaluation, and Peer Support Model Pilot)

*Teacher Practice*
The following questions are related to your **Individual Growth and Development Plan.** Please respond to the questions from your perspective as a teacher, not as a peer reviewer.

   1. Was having an IGDP useful for your development? If yes, please say more.
      Probes:
         • Did you use the rubric?
         • How did the points of contact add to your growth? If not, why not? Did you have points of contact with both peer reviewers and your summative evaluator?

2. Was it useful to complete the Self-Assessment and Peer Review Annual Summary? Why or why not?

*If teacher has had an evaluation*:

3. Tell me about your evaluation experience.
   Probes:
   - Was your IGDP a useful tool for that?
   - Did the points of contact play a role in your evaluation conversation?

4. How was your summative evaluator's use of evidence helpful for your evaluation conversation?

5. If you submitted a portfolio, how well did it generate evidence that was sufficient, meaningful, and accurate?

6. What are your overall thoughts about the Teacher Practice Component Rating?

*Looking forward:*

7. Which of the following components in the Model should be maintained for future teacher development evaluations?
   - The Performance Standards for Teacher Practice (rubric)?
   - Points of contact?
   - The Self-Assessment and Peer Review?
   - The creation of a teacher portfolio?

*Student Engagement (grades 3-12)*

1. What support did you get in understanding and interpreting your student engagement survey results? Was the support sufficient?

2. How was this process for you? Will the results of the student engagement survey be useful for your development? For your evaluation? If yes, how might you use the information?

3. Did you discuss the results of the survey with colleagues? Was that helpful? Why or why not?

4. What are your overall thoughts about the Student Engagement Component rating?

*Student Learning Goals*

1. Tell me about setting your Student Learning **<u>class goal.</u>**  Did your goal align with building or district goals?

2. How did you go about setting your mastery score?  How did that work out?

3. Did you use a different assessment for your **<u>targeted group goal</u>**? Tell me how that was developed.

4. How useful was it to have both a class goal and a targeted group goal?

5. How did you turn assessment results into a conversation with your summative evaluator?

### *Whole Model*

The following questions are related to the Whole Model.

1. How effective was the Model for developing teacher practice? To what extent are you changing your practice based on your experience with the Model (planning, instruction, environment, professionalism)?

2. What effect did the evaluation have on your students? (list of possible effects from interviews i.e. better teaching, assess student learning, determine student learning goals)

3. How well do you think the Final Performance Rating generated evidence that fairly reflected your overall performance in the summative evaluation?

4. What challenges/benefits have you experienced in your participation in the Teacher Evaluation Pilot?

5. What would you change about the MN Teacher Evaluation Pilot Model?

6. How effective was the ongoing support from the district? (may need 'Not Applicable' option)

7. Any further comments to add?

**Summer 2014 Interview Questions for MDE Staff (Tyler Livingston, Educator Effectiveness Supervisor; and Greg Keith, the Director of the Division of School Support)**

Thank you for this opportunity. Our purpose here this morning is to learn whether your vision for the pilot unfolded as you had Planned.

1. What aspects of the pilot took place as Planned and were especially successful?

2. What were the unanticipated events, issues, outcomes—positive or negative—that occurred over the course of the past year?

a. What revisions to the Model are necessary as the work shifts from pilot to full implementation?

b. How did you address unwanted or negative aspects?

3. One of the common themes in the feedback from the summative evaluators is the extreme burden on them this past year to keep on top of the many components.

a. How do you believe the summative evaluator's job could be made more manageable?

b. Are there strategies that you have gleaned from the pilot districts about helping to make the summative evaluator's job more manageable? Please say more.

4. What about districts across the state that are not in the pilot?  Are there any aspects from their teacher assessment Models that MDE might endorse or incorporate into modifications for MDE for the coming year?

5. What are your Plans for training new administrators (new districts and new arrivals in piloting districts) for fall 2014?

6. What types of MDE-sponsored training will occur this coming year?

a. Training for new administrators

b. Refresher workshops for administrators who already have had some training and experience with the MDE Model

7. How will MDE know that the teacher evaluation legislation is being implemented as proscribed in law?  Are you aware of any regulatory or monitoring actions from MDE around compliance with teacher evaluations?

8. What would you say are the most important "lessons learned" about teacher assessment Model as a result of your experiences with the pilot?

9. Did you identify any unintended consequences (positive or negative) that resulted or surfaced as a result of piloting the Model

**Value-Added Addendum**
**August 2016**


♦ **EXECUTIVE SUMMARY OF VALUE-ADDED**
**SUBCOMPONENT FINDINGS**

*Five districts participated in the value-added component of the Minnesota State Teacher Development, Evaluation, and Peer Support Model. Four of the five districts took part in the evaluation of the value-added component by sharing their experiences and perceptions in structured interviews. The districts found that participation in the value-added component underscored the importance of data use for teacher performance assessment. However, the results of data analyses summarized in value-added reports did not inform teacher learning and development per se. Of the districts who piloted the value-added component, most found that better information related to the professional development needs of teachers came from the strand scores of locally administered student assessments, such as the NWEA, in-district systematic progress monitoring in elementary reading and math, or assessments that were part of the student learning goal process.*

*There was a common concern of equity among interviewees across all districts about the value-added data. This was because only certain grade levels and certain subject areas were included in the value-added reports, and this was perceived as an issue of fairness by both administrators and teachers alike. Not all teachers in a district were being evaluated according to the same accountability measures/instruments, and this was viewed as inequitable.*

*The respondents had mixed feelings about whether or not they would continue with value-added measures. In general, they seemed very concerned about the time and effort that would be required, as well as the cost. In addition, timeliness was a concern since the value-added results received in 2015 were based on student performance in 2013-14 compared to previous school years. Some districts also mentioned delays in receiving value-added results and difficulties in accessing the web-based reports. Despite difficulties in accessing and using value-added reports, a greater awareness of the utility and importance of multiple assessment measures (including a student survey, achievement data, Individual Growth and Development Plans, peer reviewer meetings, and value-added) was an important outcome for those districts who participated in all components of the Minnesota State Teacher Development, Evaluation, and Peer Support Model.*


♦ **INTRODUCTION**

The Minnesota State Teacher Development, Evaluation, and Peer Support Model (hereafter the Model) is comprised of three components. The teacher practice component contributes 45% to a teacher's performance rating, the student engagement component contributes 20%, and the student learning and achievement component contributes 35%. The student learning and achievement component has two subcomponents, student learning goals and value-added. The Model requires all teachers to establish a measurable student learning goal. If value-added data are available, the Model requires that value-added assessment be included in the summative

evaluation of teachers who spend any portion of their instructional assignment as the teacher of record in the tested areas included in the value-added model.

Value-added modeling is a method of teacher evaluation that compares the current test performance of a teacher's students to their test performance in previous school years and to the performance of other students at the same grade level. Statistical procedures are carried out on students' past test scores to predict their future scores, based on the assumption that students will score about as well each year as they have previously. In addition to prior achievement, value-added formulas also take into account other influences on student achievement such as special education status, English learner status, and free- or reduced-price lunch eligibility. The difference between actual and predicted scores is assumed to be due to the school and to the teacher. Statisticians calculate a value-added score for each teacher by aggregating across the students in a teacher's classroom. The value-added score indicates how much an individual teacher has improved student achievement.

Five of the 17 school districts participating in the pilot of the Model chose to pilot value-added as part of the student learning and achievement component. The participating districts were: Caledonia, Granada Huntley East Chain, New Ulm, North Shore Community School, and Pine Island. The five districts used grant funds for the pilot to partner with a vendor to provide value-added assessment services. For the value-added analyses, student achievement was measured using scores on the Minnesota Comprehensive Assessments (MCAs) in reading and mathematics. MCA-Reading is administered each spring to students in grades 3-8 and grade 10, and MCA-Mathematics is administered each spring to students in grades 3-8 and grade 11. Only teachers of the grades and subjects associated with the MCA tests were included in the value-added analyses.

A number of steps needed to be accomplished up front by the school districts before the value-added analyses could be carried out and reported. The process began with roster verification where administrators matched teachers with their teaching subjects and their students. When students had multiple teachers for a particular subject, the proportion of instructional time had to be indicated for each teacher. The rosters also needed to accurately reflect any staff changes that had occurred during the year as well as student mobility. After the rosters were constructed, they were reviewed by the teachers who ensured that their classes were linked to the correct subjects, the correct students, and for the correct proportion of instructional time. After the value-added analyses were completed, the vendor prepared online reports where authorized persons could view results summarized for individual teachers, as well as results summarized at district, school, grade, and subject levels. The timeline for completion of the value-added work is shown below.

- MDE/vendor and district/vendor completed data sharing agreements as a precondition of the work.
- Districts carried out roster verification in spring 2014, and, in some cases, spring 2014 and fall 2014.
- The vendor worked with districts to make modeling and design decisions.
- Districts received MCA data on August 26, 2014.
- Districts forwarded MCA data to value-added vendor.
- The vendor carried out value-added analyses and prepared web reporting.

- The vendor scheduled and conducted training sessions with the district's summative evaluators and teachers at a time best suited for each district, which was early spring 2015 for some and late spring 2015 for others.
- The vendor communicated secure web access procedures and released value-added reports to authorized district personnel in spring 2015.
- Summative evaluators held individual meetings with teachers to discuss value-added results as part of a teacher's performance review. The schedule of evaluator/teacher meetings depended on when results were received. Some districts held meetings with teachers in spring 2015 and others in fall 2015.

It should be pointed out that many of the activities in the above timeline occurred simultaneously. In addition, there was wide variation from one district to the next with respect to when the activities were completed.

MDE formulated several questions to be addressed by the districts' experience with value added during the pilot. These questions included:

- Do the results of the value-added model inform teacher learning and development?
- Does the value-added model provide accurate and fair results for all teachers?
- What supports are needed at the individual, school, and district level to implement a value-added model for improvement planning and teacher evaluation?
- Do individual districts have the capacity and resources to develop, interpret, and report value-added data?

## ♦ DATA COLLECTION

In the final phase of its evaluation of the Model, the Center for Applied Research and Educational Improvement (CAREI) gathered data that would address questions concerning the value-added assessment of teachers. CAREI contacted 13 administrators from the five districts that had piloted the value-added assessment to invite them to participate in an interview. The administrators were either district superintendents or school principals. Initial invitations were sent via email near the end of the 2014-15 school year and non-respondents were contacted a second time by email and then by phone. Four of the 13 administrators were not able to answer questions about value-added assessments in their districts either because they had not been directly involved with the value-added processes or had left the district and were no longer available to provide comments. Three other administrators had very recently received a link to their value-added reports and had not yet reviewed them. Therefore, in June 2015, CAREI evaluators were only able to conduct phone interviews with three administrators regarding their experience with value added. Six administrators who had not been interviewed in June were contacted again in the fall of 2015, and two of these individuals agreed to be interviewed. Overall, interviews were completed with a total of five administrators (three in June and two in October) who represented four of the five districts that participated in the value-added pilot.

The interview protocol was developed collaboratively by CAREI and MDE staff. The questions concerned the upfront work required by the districts, the interpretation of the results, the

consistency of value-added results with other measures of a teacher's performance, how the results would be used, the benefits of using value-added, the challenges associated with value-added assessment, and the district's plans to continue the use of value-added in the future. A copy of the interview protocol is provided in the Value-Added Addendum Appendix.

# ♦ RESULTS

The interview protocol (see Value-Added Addendum Appendix) is used in the following sections to organize the responses from the five interviewees, all of whom were summative evaluators. To help the reader understand the interviewees' perspectives, verbatim responses are included in the results.

### *Upfront work to get value-added reports*

Only four of the five interviewees had been directly involved with the roster verification process and other upfront work and were able to comment on it. These four administrators were from three different districts.

- An administrator from one of the districts praised the assistance given by the vendor and stated that the vendor walked them through the process and provided clear instructions that were easy to follow. This administrator felt that everything went very smoothly.

- An administrator from another district talked about a need for better communication with the vendor. For example, the administrator said that they had placed the required teacher roster data into a safe online place as instructed, but they were still waiting to hear back from the vendor with a password [at the time of this interview]. Another administrator in this district who was interviewed corroborated the great amount of difficulty in accessing the data: "I encountered a lot of difficulties, and I have to tell you that I never actually got access… The superintendent sent me the email with the link and who to contact. I emailed that contact and attached the letter for me to have access. The person answered me, 'Sure thing, I'll make sure to get you access on [date]', but then I never got access. I did email one more time after that and still haven't heard back."

- The superintendent of the third district had to do the sharing of the value-added results (downloading reports, etc.) because the principal was not allowed access by the vendor, who had "closed the vault" [this was the vendor's explanation for why the principal could not get access]. Furthermore, the superintendent viewed the roster verification process as a lot of work that ultimately did not result in useful information.

### *Evaluating teacher performance – consistent with other results?*

All five interviewees had strong views on whether or not the value-added reports provided useful information for the assessment of teacher performance. The value-added reports were not received by the districts until spring 2015 at the earliest. This, in part, explains the perception of data being "old," since to calculate value-added scores, student performance in the 2013-14

school year was compared to performance in previous school years. The analyses were carried out during the 2014-15 school year, and the results were not viewed by some summative evaluators until the beginning of the 2015-16 school year. The comments from the interviewees reveal how each district was able to use or not use the value-added data and reports in the evaluation of teachers.

- "I've been through many different presentations on this, and nobody has been able to really speak to how you can measure value-added for each and every teacher with any sense of validity, because there are so many other variables that contribute to student achievement." Also, this evaluator noted that, since the value-added reports were not available for more than a year, the data were "old" and the students had moved on at least two grade levels. Thus, the reports were not seen as providing useful information for evaluating teachers in the current year.

- When asked about the reports showing the red and green cells for individual teachers, and how the evaluator decided whether performance levels were due to the teacher, to the curriculum, a staff transition, or maybe something else, the reply was, "That was where we ran into a lot of debate…. I conferred with [a colleague] and we would look at a teacher's performance… with our observations and other results. We said 'this doesn't match up.' That's when we realized that you take the group of students that the teacher had and you follow them in the following year. It turns out that class was just low performing." In another case, where a teacher had already been perceived as being weak, as based on other evaluative measures, the value-added report was confirmatory, and strengthened the decision to reassign the teacher.[a]

- In this district, because the value-added results came after the school year was over, the summative evaluator did not complete the summative meetings with either elementary or middle school teachers until late fall 2015. This evaluator was skeptical about the fairness of the results—"A good teacher who kept the 'low' kids got a yellow indicator and looked bad. She had some of the toughest kids. This felt unfair."[a]

- Not having the value-added results did not negatively affect this summative evaluator's ability to provide a comprehensive performance evaluation. This was because the district also has NWEA scores for many years for the students: "Over the years for those teachers [who were not effective] with the kids, I could see that the achievement level was down." Also, this evaluator found consistency with other measures of teacher performance, such as the student survey and student achievement. "When I look at the student surveys and I look particularly at how I view a teacher's ability to develop relationships, it's very accurate. The students are assessing the teachers the same way that I am. I can see a lot of relationship building, and the students feel it. …I felt when the teachers weren't developing good relationships, then the students were not achieving as well as other students."

- This respondent was unable to comment on the consistency between value-added and other teacher performance assessments, as the respondent had not yet [at the time of the interview] been able to view the value-added results.

*Were value-added reports useful for PD decisions? For evaluating the effectiveness of programs/initiatives?*

In general, the five respondents saw limited usefulness in value-added reports for making decisions regarding PD or for evaluating the effectiveness of programs and initiatives.

- "The information we got back was not worth it. All teachers looked pretty average and pretty generic. Even with a new curriculum, we could not see any specific indications of what parts were working well and which needed more PD."

- "We really have no more information than before to move forward with… There were no strand scores in reading. There weren't any subsections of which skills in reading for which the kids needed more help, for example, in vocabulary or in phonics. This is not useful information for targeting PD or improving your teaching."

- This respondent indicated that the district will continue to base decisions about district-wide PD on RTI strand scores for reading and math. "When we study that data, we do try to develop professional development around some of the needs that we may see. Like we may focus one year on a particular reading strategy or if we see a strand where we are low in math over three years, we may pick up on that strand and what strategies can we give our teachers to improve their instruction on that strand."

- In this district, the summative evaluator used the end-of-the-year results of the Individual Growth and Development Plan (IGDP) to decide whether or not a teacher needed professional development in a particular area. "Still, that's an area where we need improvement. Some [teachers' professional growth goals] were excellent and wonderful, but there's still a few out there that just don't understand that this is about your professional development, your growth. … Some were misunderstanding what it's all about. It's not about changing parents. It's not about changing the kids. It's about changing you" [emphasis provided by interviewee].

- This summative evaluator felt that the value-added reports were useful for teacher PD that occurs within PLCs because the emphasis should be on overall results rather than results for individual teachers. "Yes, the findings will be used as a focal point for PLCs, as we only let the teachers look at their own grade level, and not at their individual reports. …It's proof of the direction we need to go. Sometimes we get kind of set in our ways and we think we are doing OK because our scores are not that bad. But yet when you look at the overall results, the added value, and then all of a sudden you start thinking we are not doing as well as we thought we were. I mean the scores are not that terrible, yet this is the difference factor. And I think that we really need to look at the whole picture."

In addition to comments concerning usefulness of value-added for making decisions regarding professional development, programs, and initiatives, two interviewees also commented on the usefulness of "old" value-added data. One respondent said, "It is such old data, it is hard to have it feel real. The data is from three classrooms ago, and most teachers could not readily recall the

needs of the kids in those classes." The other respondent said, "The data is from two or three years ago, so I don't think we're going to be spending a lot of time with it."

*Benefits of value-added reports*

Two of the interviewees were not able to identify benefits associated with value-added reports because they had not yet seen the value-added reports for their districts. The benefits mentioned by the other three interviewees included being able to make comparisons between grade levels, encouraging administrators to focus on expected learning gains, and serving as a "pat on the back" for teachers whose students made greater progress than expected.

- By seeing value-added reports for each grade, "…as a district we could see that there were a couple of problems with some grade levels. …we saw how fourth grade was pretty good and then we saw a huge drop in the district. And sixth grade drove the math up."

- In this district, the value-added reports were considered beneficial because the color codes encouraged them to focus on learning gains. "It was absolutely mind boggling when we first saw the results and we saw the different colors. It was not what I expected at all. And it was, I mean, it was in there in color and you can't miss it. It's just so obvious. …when you look at some of the red numbers and the lack of growth, obviously we need to have something for the teachers to set for a goal."[a]

- "Most teachers felt OK with their scores. Those that got a green indicator felt they got a 'pat on the back' with the data."[a]

*Weaknesses/challenges with value-added measures*

All interviewees pointed out weaknesses and challenges with value-added measures, with some identifying two or more. The weaknesses and challenges mentioned by the interviewees included value-added results that were inconsistent with other teacher performance indicators, cumbersome confidentiality procedures, difficulty getting results, results received too late to be meaningful for teacher evaluation, lack of teacher buy-in, results pertaining only to teachers who teach math and reading, results pertaining only to teachers of third grade and above, and results placing too much emphasis on a single numerical outcome.

- This summative evaluator questioned the accuracy of value-added measures for teacher evaluation purposes. The respondent said, "It makes you almost ask yourself, 'Am I evaluating this teacher in the proper way or am I missing something that I can't catch?' Because the results just aren't adding up."

- This summative evaluator said that "the superintendent believes the ideas behind having value-added data are great but felt that there was so much confidentiality or worries about having people log in to safe places, that it slowed the progress."

- The value-added reports were never obtained in this district, as the "data vault was closed" [per the reply from the website]. "I had multiple people try and open the webpage links that I was sent and we can't get that open either." At the time of the interview, the contact person for the vendor had not yet responded back with a new password to enter the data site.

- Commenting on the lateness of the results, this summative evaluator said, "We didn't get the results until quite late and it was well past the time that the summative evaluation would have taken place."

- This evaluator reported that the webinars were not clear for the teachers and teachers were expected to "trust the mystical statistical process." As a result, the teachers "did not really buy in" to the value-added component. "It was a big deal, a lot of work, to get the value-added report, and it was not useful. It certainly did not change practice." …"I do not have a lot of confidence in the process."

- This evaluator was concerned about value-added measures applying only to teachers of math and reading, the subjects tested by MCAs. The evaluator said, "The biggest struggle is, you know, equity in terms of all staff being able to be measured. It seems like there's only certain teachers that can be measured."

- Similarly, this evaluator was concerned that value-added did not apply to lower grades. "One of the weaknesses is that we don't have data for the early grades to use it. …Not that I want to have more assessment, but it would be nice to know even earlier than third grade where kids are at."

- This evaluator questioned the appropriateness of assigning a numerical score to teacher performance. "With value-added, I feel like I am grading someone. And I would prefer to get away from grading people. Although I recognize their weaknesses and I can work with them, I don't need to put a number on it or a grade on it. And, I don't like that when people go around and compare themselves with other teachers because of that number. There's a lot of interpretation that can go into that number. And just because somebody got a number and maybe was stellar on one day doesn't mean they're stellar all year long. So I would rather have conversations with people and have them reflect and talk about what they do well and talk about where they feel they need to work. … The number doesn't do anything but cause hurt feelings."

*Suggestions/recommendations*

Interviewees from three districts provided suggestions for improving various aspects of the value-added process. The suggestions concerned improving access to the value-added reports, providing training to all summative evaluators, and making it possible to make district-to-district comparison of value-added results.

- "For instance, with the email with the password we got from the company— why don't they just send us the school and teacher reports? So, it would just be sending [the value-

added reports] to the superintendent. It's not like it's sending them to everybody, and if we had access to that, then we could have used that much more easily."

- "[The district] selected half of the administrators for training [in the value-added process] and half were not trained…. However, [I believe] each summative evaluator doing an observation should have been involved." This summative evaluator did not attend the training, and, consequently, did not use any of the value-added results to evaluate teacher performance.

- Previously, "over different years, we would compare ourselves to other school districts. But we weren't given that opportunity with [our value-added provider]. I don't know if that is possible or not."

### *Continue or not continue with value-added measures?*

The respondents had mixed feelings about whether or not they would continue with value-added measures. In general, they seemed very concerned about the time and effort that would be required and the cost. Also mentioned were delays in getting the reports and concerns that value-added does not apply to all the teachers in the district.

- "I think that that is something that we would want to do. ...As much of a pain as it was to get everything going, I think the results are very useful."

- I think [value-added] is an idea that sounds interesting, but if it's going to cost above and beyond what we're already doing for TD and E (Teacher Development and Evaluation), I don't think that it is something that many district schools will pursue…I don't think we will."

- "The headache that it takes to get the reports is not worth it. The [two kinds of] 'costs' are prohibitive. One, it takes human resources time to download the reports, and two, the time it takes for principals and teachers to do this."

- "Next year we are not going to be using [a value-added vendor]. I know that. It's too expensive. … The reports are fascinating to look at. They were…. But it's a whole lot of work for us….And if it weren't [so much work], it would be valuable information to have."

- "I think it [value-added report] could be useful, but the problem is there are so many flaws in the process… [I am concerned about] the delays in getting the reports and the fact that it feels inequitable because it does not involve all staff, and only certain teachers are able to be measured."

*Additional comments*

When asked about the long-term benefit of value-added assessment, a summative evaluator had this to say: "You know, the best way that schools can do that is something that we have been doing for about year, is actually what the NWEA (MAP) assessment does for us, especially in an elementary school where you might have four or five sections of a grade level. And if you have the same four or five teachers, you can see the teachers that get the most growth from students from year to year. That is a really good indicator of value-added…. With NWEA you get the scores back in a timely fashion, so you know within that school year how the teacher is doing and how the kids are doing."

Also commenting on data that are already being collected, another summative evaluator said, "Our weekly data collection with progress monitoring tools is, by far, the best strategy for improving instruction. The data from the MCAs, the general overall data we get for reading and math, is not helpful."

## ♦ SUMMARY AND CONCLUSIONS

MDE formulated several questions to be addressed by the districts' experience with value-added measures during the pilot. These questions included:

- Do the results of the value-added model inform teacher learning and development?
- Does the value-added model provide accurate and fair results for all teachers?
- What supports are needed at the individual, school, and district level to implement a value-added model for improvement planning and teacher evaluation?
- Do individual districts have the capacity and resources to develop, interpret, and report value-added data?

The answers to the questions above can be summarized as follows:

In general, the results from the data in value-added reports did not inform teacher learning and development per se. Of the districts who piloted the value-added component, most found that better information related to the professional development needs of teachers came from the strand scores of locally administered student assessments, such as the NWEA or systematic progress monitoring. Just as it is for the MCA data, the value-added results do not indicate the specific areas of reading and math that are lowest and, therefore, the strategic focus for teacher development is missing. However, it is noteworthy that one district in this study will be using the value-added data to inform the direction and focus for their PLCs in the next year.

There was a common concern of equity among interviewees across all districts about the value-added data. This is because only certain grade levels and certain subject areas were included in the value-added reports, and this is perceived as an issue of fairness. Not all teachers in a district are being evaluated according to the same accountability measures/instruments, and this was viewed as inequitable.

The supports needed to implement a value-added model at all levels center around issues of time and money. A great deal of time was needed to create and verify rosters of teachers and their students, for which some districts used the grant funds they received from MDE to pilot the Model to buy time of a person to do this. Without continued financial support, all of the districts involved in this component felt that the costs were prohibitive relative to the benefits of having a value-added report. Also, the very large time lag between when a school year was over and when the value-added report became available to summative evaluators (a year or more for some), caused the immediate benefit of having a value-added report to become moot.

All of the districts who participated in this interview process noted that they still had great needs in developing the local capacity to better interpret and utilize data. Both deep knowledge and sufficient time are resources that the districts do not have in order to adequately use what might be available to them via value-added measures.

In summary, the districts found that participation in the value-added component underscored the importance of data use for teacher performance assessment. This was true even for the districts that had not yet received their value-added results. Bringing greater awareness of the utility and importance of multiple assessment measures (student survey, achievement data, IGDPs, peer reviewer meetings, etc.) was an important outcome for districts who participated in all components of the Minnesota State Teacher Development, Evaluation, and Peer Support Model. To that end, the many components in the Model revealed that, as with students, a measure of a teacher's performance is a compilation of a variety of forms of assessment. With that greater awareness should, hopefully, come greater learning in Minnesota classrooms.

---

[a] Note: The different colors in the reports indicated if learning gains were above expectations (green), at expectations (yellow), or below expectations (red).

# Appendix:  Value-Added Interview Protocol

1. First, tell me about the upfront work that you had to do in order to get value-added reports. Was this a fairly easy process or did you encounter difficulties? If you ran into difficulties, what were they?

2. For evaluating a teacher's performance, did you find the value-added results to be consistent with other measures you had such as the teacher's Student Learning Goals (SLG), Individual Growth and Development Plan (IGDP), or the student survey results? Can you give me some examples of consistencies or inconsistencies that you observed?

3. How did you go about deciding whether strong results—the green cells—or weak results—the red cells—were probably due to curriculum, staff transitions, teacher effectiveness, or something else?

4. What about the content of the value-added reports? How did you use or how will you use the reports to make decisions about:
   a. Professional development for teachers?
   b. The effectiveness of programs or initiatives that your school is implementing?

5. Overall, what would you say are some of the greatest benefits or positive outcomes associated with having the value-added reports?

6. Overall, what would you say were some of the greatest weaknesses or challenges associated with the value-added reports?

7. Based on your experience, do you have any suggestions or recommendations regarding the use of value-added measures?

8. What are your thoughts about having your district continuing value-added measures next school year and subsequent years? Probe: Does your district have the capacity and resources to develop, interpret, and report value-added data to continue its use?

9. Do you have any further comments that you think would be useful for MDE or other policy makers to know about the value-added component?