

Background and Analysis Methods

To support ongoing professional growth for all educators, the Colorado Department of Education (CDE) has developed a model evaluation system that districts may use to evaluate their teachers, which meets all requirements of Senate Bill 10-191 and State Board of Education Rules. The Colorado State Model Evaluation System (or state model system) for teachers was piloted in 26 districts during the 2012-13 school year and is an optional evaluation system used by 160 Colorado districts in 2013-14 school year. Feedback was collected from the initial 26 pilot districts (through surveys, interviews and focus groups) to inform improvements to the evaluation system.

In the fall of 2012, pilot teachers completed an initial survey to establish a baseline understanding of their opinions of their current evaluation system (prior to piloting the state model system). Then, they were introduced to and evaluated by the state model system over the course of the 2012-13 school year. Teachers and principals were surveyed at the end of the 2012-13 school year to gauge their perceptions of the state model system for teachers.

This report expands on the [2012-13 Teacher System Pilot Baseline and Feedback Survey Data report](#), which showed the initial findings and gave an overview of the survey results. This new report further analyzes the survey data in order to give more detailed information about survey results, including responses on additional items and comparisons of teacher and principal responses. The survey was completed by 749 teachers and 73 principals. The combined response rate was 43 percent.

A factor analysis was run to group the teacher and principal responses according to the strongest relationships among responses. This enabled the identification of general topic areas which were underlying the survey questions. These topic areas were defined by what the grouped questions shared in common. The results are reported both by these topic areas (i.e. as aggregate variables) and also in separate charts as individual questions within these topics.

Teachers and principals were compared by how they answered the survey questions. The principals were in stronger agreement than the teachers in three out of the four areas in which they can be compared (note that the statistically significant differences are indicated by asterisks next to the topic area name).

Finally, the report concludes with an appendix on qualitative findings from the survey's open-ended questions.

Findings

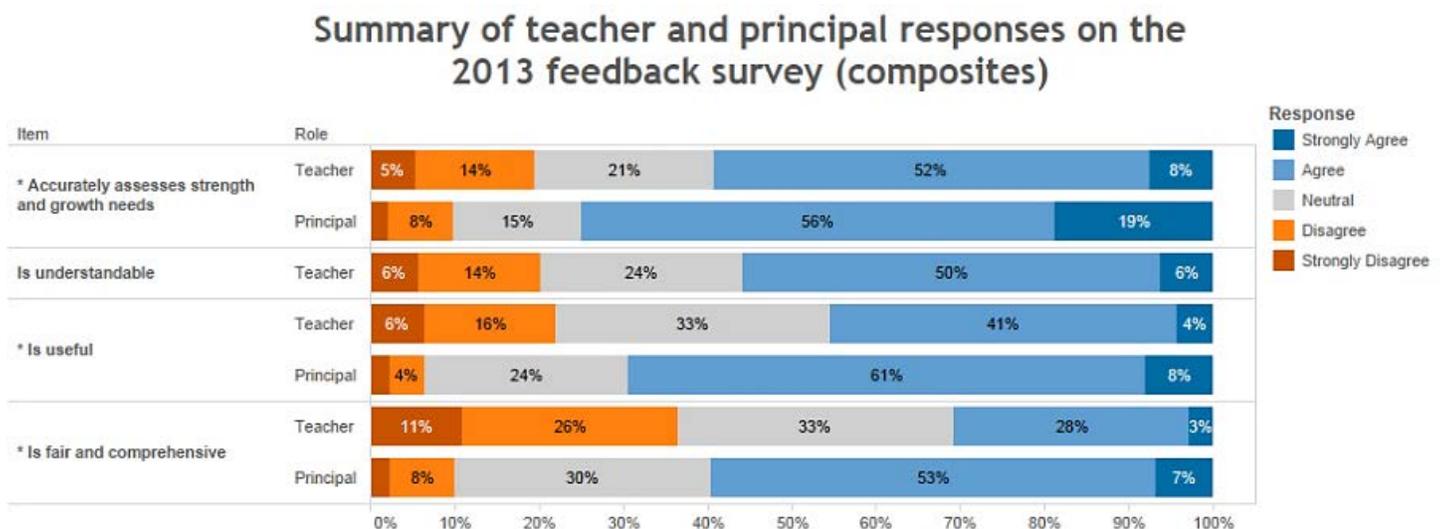
Initial factor analyses revealed six topic areas which the survey questions could be grouped. The topics are: *Accurately assessing strength and growth needs, understanding the system and how to get information, usefulness of the system, being fair and comprehensive, high time burden, and worth the time using the system.*

The response key and corresponding colors represented in Exhibit 1 show the five possible responses for survey questions, which include strongly disagree, disagree, neutral, agree, and strongly agree. Exhibit 1 represents an aggregation of question results for four of the six question topics. The figures shown were calculated by averaging the percent of responses that fell within each of these five response categories for the questions within each topic (The high time burden topic and the worth the time using the system are separated out because of a different question format and response options). The stacked bar chart in Exhibit 1 below (and all others in the report) is ordered from most positive teacher ratings to least positive teacher ratings.

In the four topic areas in Exhibit 1, there were two topic areas in which the teachers were above 50 percent agreement. *Accurately assessing strength and growth needs* is the topic in which teachers and principals were most favorable to the state model system. The teacher median agreement percent (combining agree and strongly agree) for the seven questions which make up that topic was 60 percent, with four questions at 70 percent agreement or higher in this topic area. The other topic area in which teachers were above 50 percent agreement was on *understanding the system and how to get information*, on which 56 percent of teachers agreed.

The last two sets of bars on Exhibit 1, shows that less than half of the teachers agreed that the state model system is *useful* and *is fair and comprehensive*. However, it is also informative to consider the percentage of teachers who gave neutral responses because they weren't explicitly dissatisfied with the state model system. When the neutral responses were added to the positive, 78 percent of teachers were in this larger group for *system is useful*, and about 64 percent for the *fair and comprehensive* topic questions. Since this data is from the first year of implementation, it's not surprising that there is still learning to do on the system and more training is needed to ensure that all educators feel that the system is fair and comprehensive. In the first year of any large-scale change effort, one would expect these types of perceptions.

Exhibit 1. Teacher and principal percent agreement on feedback survey topic areas

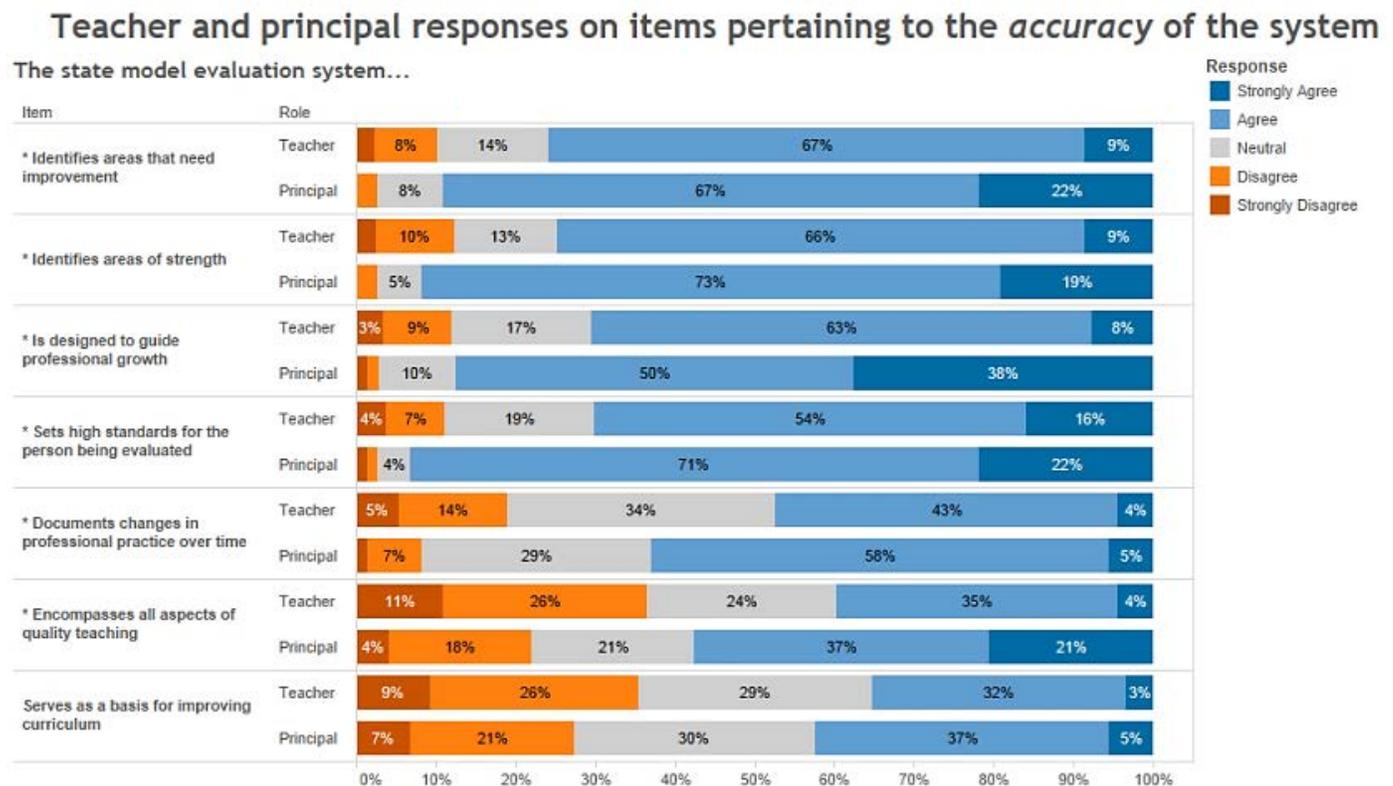


* = Statistically significant difference between teacher and principal responses.

While Exhibit 1 takes an overall view of the large topic areas, Exhibits 2 through 5 show the details of the topic questions, that is, the teacher and principal percent positive responses to the *individual* questions on the survey. Remember that the survey was completed by 749 teachers and 73 principals, so the teachers outnumber the principals about 10 to one. Even with the relatively small number of principals in the survey, the differences in responses were still found to be most often statistically significant, with principals responding more positively. Principals might respond more positively because they have piloted the system for an additional year, thus they have more experience with the system. Since principals are also the evaluators as opposed to the person being evaluated, they may have more positive perceptions.

Exhibit 2, pertaining to the accuracy of assessing strength and growth needs and curriculum improvement needs, shows very positive results for the first four questions on identifying areas for improvement and areas of strength, guiding professional growth, and setting high standards. The agreement results for both principals and teachers are quite high, at or above 70 percent for teachers and well above 80 percent for principals. These areas represent some of the major goals of the state model system. The lowest positive response for both teachers and principals (at around 40 percent for each) was for the last question at the bottom of this chart which is “serves as a basis for improving curriculum.” But when neutrals are added to the agreement group, 64 percent of teachers are in this agree or neutral group.

Exhibit 2. Accuracy of the state model system’s assessment of strength and growth needs



* = Statistically significant difference between teacher and principal responses.

Exhibit 3, understanding the system and knowing how to get information on the system, shows results from teachers only since principals were not asked questions on this topic. In the aggregate, more than half of teachers indicated that they had good understanding of the system or know where to go to get questions answered. They responded most positively to direct questions about understanding the system and knowing who to go to if they have questions.

Concerning the last question, *the information I received from my district was helpful in understanding the evaluation system*, 43 percent stated positively that information that they received from their district leaders was helpful for them to understand the state model system.

Exhibit 3. Understanding the state model system

Teacher (only) responses on items pertaining to their understanding of the system

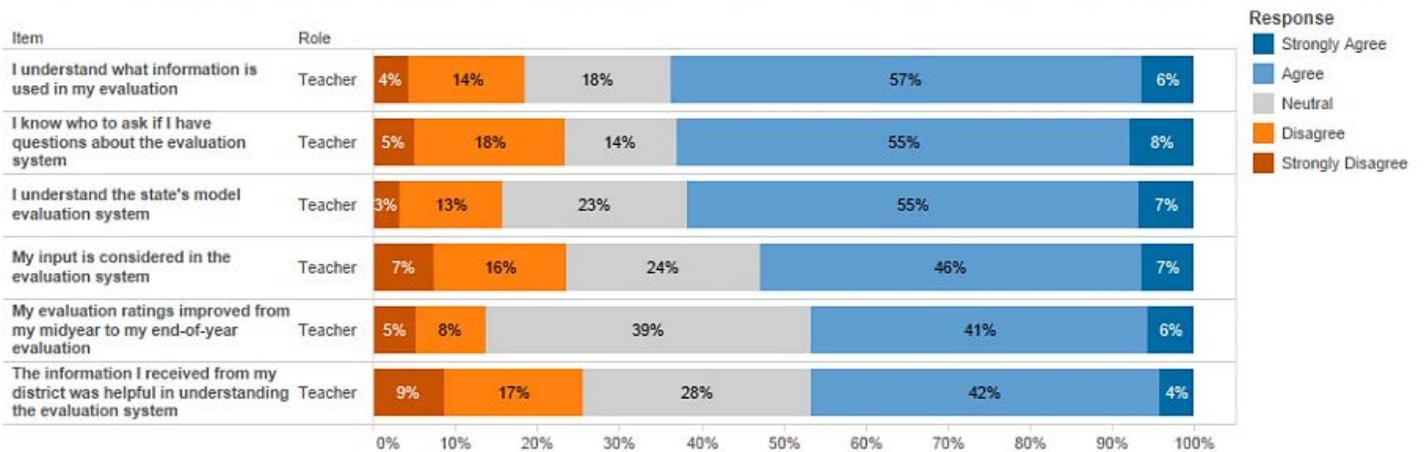
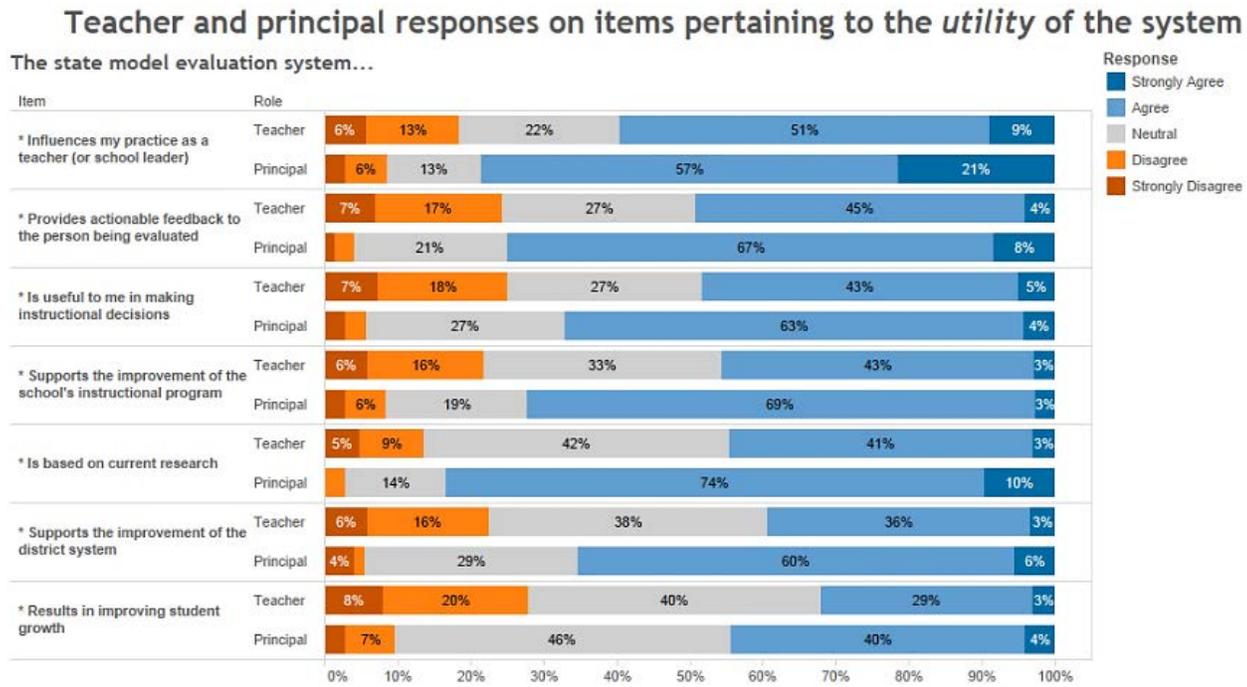


Exhibit 4 represents one of the largest topic groups of questions in the survey, all pertaining to the usefulness of the state model system. The most positive response was in regard to the question as to whether the system *influences my practice as a teacher or school leader*, for which 60 percent of teachers and 78 percent of principals answered positively.

The least positive response was to the question pertaining to the system’s impact on student growth. Thirty-two percent of teachers responded positively to the question concerning whether the system *results in improving student growth*. However, when the “neutral” category is added it jumped to 72 percent. In this first year of implementation, one would expect these numbers as educators might be withholding judgment until they have an adequate chance to see student results.

Exhibit 4. Usefulness of the system



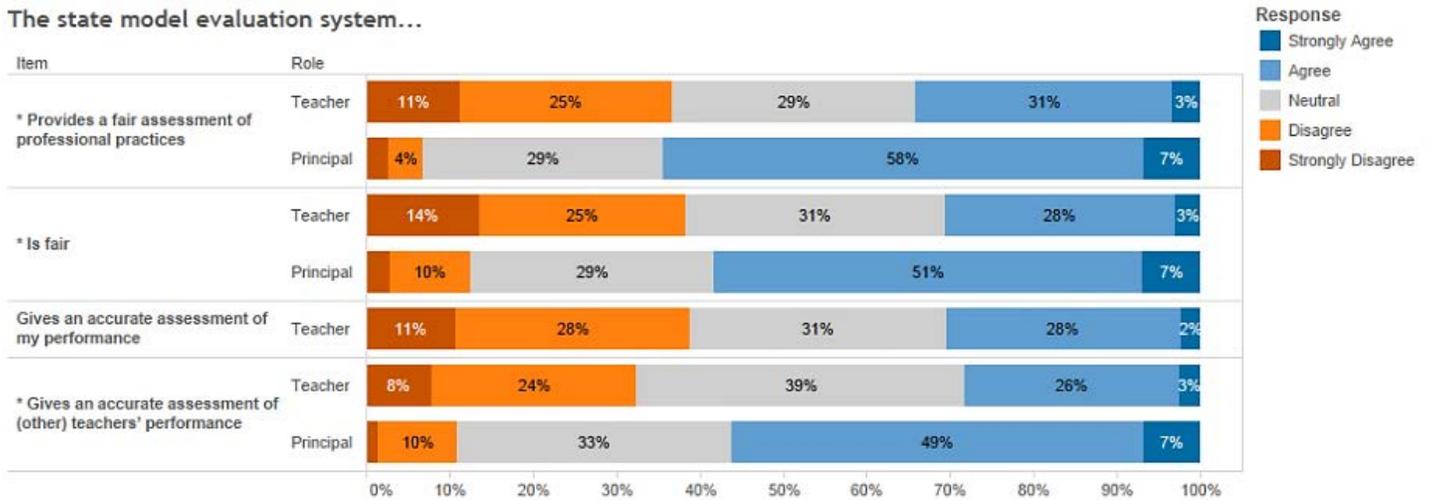
* = Statistically significant difference between teacher and principal responses.

Exhibit 5 reports on findings on whether the system was *fair and comprehensive*. The average of teacher positive responses for this question set was 31 percent, while principals were 60 percent positive. Principals were much more likely to agree that the state model system is fair and gives a full and accurate assessment of teachers' performance. This data most likely reflects the variety of ways that districts tried to implement the system in the first year. Fidelity of implementation across the state was varied and district administrators were learning a whole new process. This likely had impact on the consistency of evaluation procedures both within and between districts.

Exhibit 5. A fair and comprehensive system

Teacher and principal responses on items pertaining to the *fairness* of the system

The state model evaluation system...

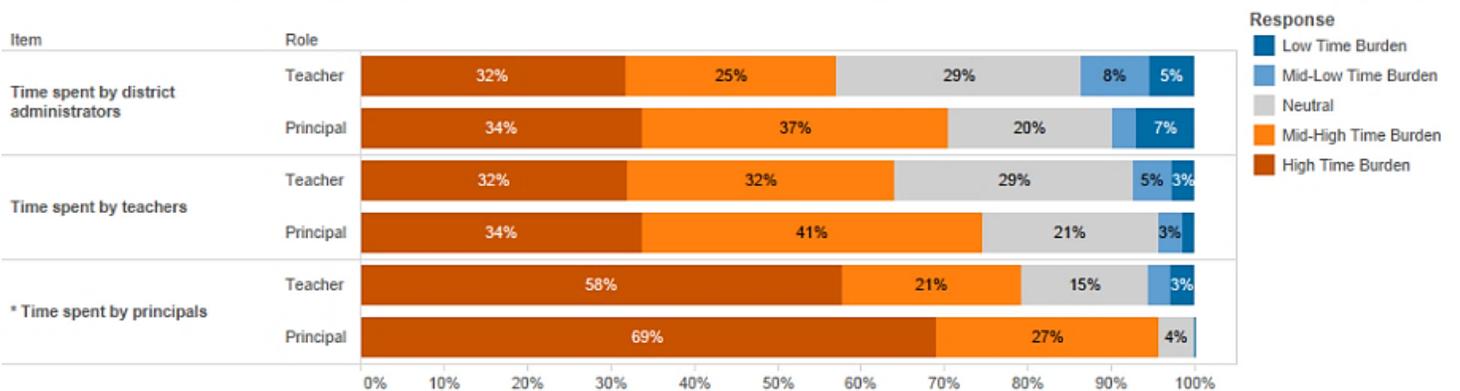


* = Statistically significant difference between teacher and principal responses.

Switching gears to focus on the time spent on the state model system, Exhibit 6 indicates from both the principal and teacher perspective that a lot of time is being spent implementing the system. Both groups indicate that district administrators, teachers and especially principals experience a time burden. The responses to the question on principal time spent were quite high for both principals and teachers, with 79 percent of teachers and 96 percent of principals entering on the *high time burden* side for this question.

Exhibit 6. Time burden of the system

Teacher and principal responses on items pertaining to the *time burden* of the system



* = Statistically significant difference between teacher and principal responses.

Considering where educators are spending time on the system, teachers and principals were asked to estimate the time that it took for them to accomplish specific components of the state model teacher evaluation system. The following eight tables (7a to 7h) show their responses to this request. The first table reflects teachers' times, while the remaining seven represent principal times.

Exhibit 7a. Time for teachers to finish components of an evaluation

Survey item	Time taken by teachers	Number (and %) of teachers
Estimate the amount of time it took to finish your components of your evaluation this year (e.g., self-assessment, goal-setting, etc.).	0 - 30 minutes	28 (4%)
	31 - 60 minutes (1/2 hour - 1 hour)	123 (17%)
	61 - 120 minutes (1-2 hours)	240 (34%)
	121 - 240 minutes (2-4 hours)	221 (31%)
	More than 4 hours	104 (15%)

Exhibit 7b. Time for principals spent in training or training others on state model system

Survey item	Time taken by principals	Number (and %) of principals
Estimate the amount of time amount of time spent in training or training others on the state model system.	0 to 4 hours	5 (14%)
	4.5 to 8 hours	13 (35%)
	8.5 to 16 hours	8 (22%)
	More than 16 hours	11 (30%)
	I did not participate in any training	0 (0%)

Exhibit 7c. Time for principals to complete the review of each teacher's goals and/or self-assessment

Survey item	Time taken by principals	Number (and %) of principals
Estimate the amount of time it took for you to complete the review of each teacher's professional goals and/or self-assessment	0 - 30 minutes	3 (8%)
	31 - 60 minutes (1/2 hour - 1 hour)	12 (32%)
	61 - 120 minutes (1-2 hours)	13 (35%)
	121 - 240 minutes (2-4 hours)	9 (24%)
	More than 4 hours	0 (0%)
	I did not complete this portion of the evaluation process	0 (0%)

Exhibit 7d. Time for principals to complete the paperwork associated with each teacher's evaluation

Survey item	Time taken by principals	Number (and %) of principals
Estimate the amount of time it took for you to complete the paperwork associated with each teacher's evaluation.	0 - 30 minutes	0 (0%)
	31 - 60 minutes (1/2 hour - 1 hour)	5 (14%)
	61 - 120 minutes (1-2 hours)	11 (30%)
	121 - 240 minutes (2-4 hours)	15 (41%)
	More than 4 hours	5 (14%)
	I did not complete this portion of the evaluation process	1 (3%)

Exhibit 7e. Time for principals to fill out the mid-year rubric for each teacher

Survey item	Time taken by principals	Number (and %) of principals
For the mid-year conversation, estimate the amount of time it took for you to fill out the rubric for each teacher.	0 - 30 minutes	5 (14%)
	31 - 60 minutes (1/2 hour - 1 hour)	13 (35%)
	61 - 120 minutes (1-2 hours)	12 (32%)
	121 - 240 minutes (2-4 hours)	1 (3%)
	More than 4 hours	0 (0%)
	I did not complete this portion of the evaluation process	6 (16%)

Exhibit 7f. Time for principals to complete the mid-year feedback conversation with each teacher

Survey item	Time taken by principals	Number (and %) of principals
For the mid-year conversation, estimate the amount of time it took for you to complete the feedback conversation with each teacher.	0 - 30 minutes	7 (19%)
	31 - 60 minutes (1/2 hour - 1 hour)	19 (51%)
	61 - 120 minutes (1-2 hours)	7 (19%)
	121 - 240 minutes (2-4 hours)	1 (3%)
	More than 4 hours	0 (0%)
	I did not complete this portion of the evaluation process	3 (8%)

Exhibit 7g. Time for principals to fill out the end-of-year rubric for each teacher

Survey item	Time taken by principals	Number (and %) of principals
For the end-of-year conversation, estimate the amount of time it took for you to fill out the rubric for each teacher.	0 - 30 minutes	4 (11%)
	31 - 60 minutes (1/2 hour - 1 hour)	18 (49%)
	61 - 120 minutes (1-2 hours)	10 (27%)
	121 - 240 minutes (2-4 hours)	3 (8%)
	More than 4 hours	1 (3%)
	I did not complete this portion of the evaluation process	1 (3%)

Exhibit 7h. Time for principals to complete the end-of-year feedback conversation with each teacher

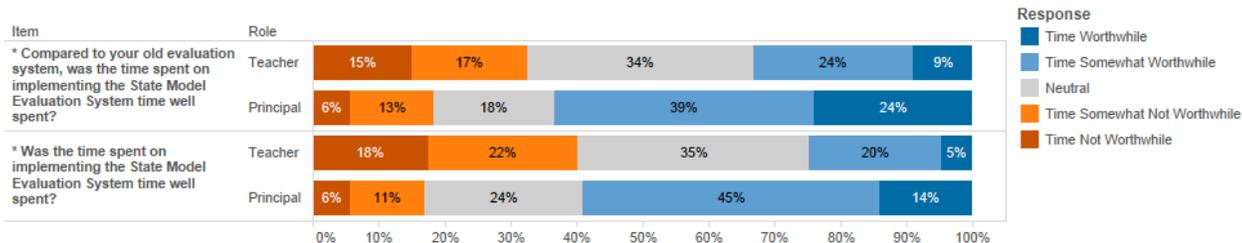
Survey item	Time taken by principals	Number (and %) of principals
For the end-of-year conversation, estimate the amount of time it took for you to complete the feedback conversation with each teacher.	0 - 30 minutes	5 (14%)
	31 - 60 minutes (1/2 hour - 1 hour)	23 (62%)
	61 - 120 minutes (1-2 hours)	5 (14%)
	121 - 240 minutes (2-4 hours)	1 (3%)
	More than 4 hours	0 (0%)
	I did not complete this portion of the evaluation process	3 (8%)

The shaded rows in Exhibits 7a-h represent the most common amounts of the time taken by principals for each of these survey items. If these are tallied (represented by these shaded areas), the time spent by these principals on the whole system for each teacher ranges from 5.2 to 10.1 hours.

Exhibit 8 shows responses to questions about whether the time using the system is well spent. In the first pair of bars, they are asked to compare the system to the system that they were using before the state model system was implemented. More than 60 percent of teachers either agreed or were neutral that the new system was time well spent compared to their old evaluation system.

Exhibit 8. Worth the time

Teacher and principal responses on items pertaining to whether the system is worth the time

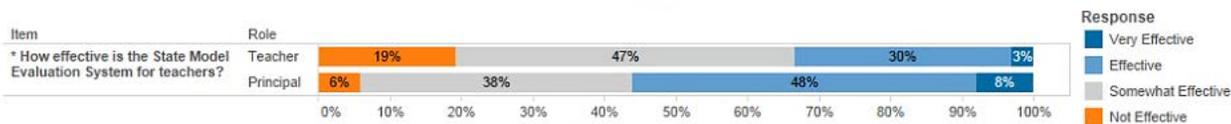


* = Statistically significant difference between teacher and principal responses.

Exhibit 9 shows the responses to the question “How effective is the State Model System for teachers?” The system was considered to be at least somewhat effective by about 81 percent of teachers (note that *somewhat effective* is something like a neutral response) and 94 percent of principals.

Exhibit 9. The effectiveness of the state model system

Teacher and principal responses on items pertaining to the effectiveness of the system



* = Statistically significant difference between teacher and principal responses.

Conclusion

The overall sense from the quantitative analysis of the survey was that the majority of principals were quite favorable in their perceptions of the state model system. Teachers were generally favorable or neutral on the system for survey questions pertaining to the accuracy of the system, understanding of the system, utility of the system, and whether the system is worth the time. There were two areas of concern for educators. First, both principals and teachers saw the system as requiring a large amount of time. The other area where there was less positive agreement among teachers was the perceived fairness of the system.

The high time burden should be alleviated as principals and teachers become more familiar with the state model system, as other states have reported with similar systems. The fairness of the system will likely depend on the principals and teachers becoming more comfortable with the process and developing more uniform evaluation processes. Studies are ongoing to validate districts’ implementation of the state model system and to look for ways to improve the comparability among teacher ratings, thereby improving the fairness of the system.

CDE has used this survey data to inform changes to the system, including shortening the rubric and providing additional resources to explain components and processes in the system. CDE has added sections to its training to address some of these concerns such as how to give more actionable feedback to educators. In addition, CDE is building supports such as Elevate Colorado (an inter-rater agreement system) and a [Resource Guide for Deepening the Understanding of Teacher's Professional Practices](#) to help explain practices and support fair and consistent interpretations of practice.

This report shows that there is still a way to go in realizing the full intent of the evaluation system, but there is reason to believe that elements of implementation can be improved. As a first set of data on perceptions about the system, it is not surprising to see some evidence of trepidation. CDE will be working with districts to improve these perceptions over time through support, training and technical assistance on the implementation of their evaluation systems.

Appendix: Analysis of Teacher and Principal Open-ended Responses

This analysis was conducted by Julia Kantor, University of Colorado at Boulder

Survey purpose and goals

In the spring of 2013, CDE administered a survey to educators in the 26 districts that had spent the 2012-2013 school year piloting the state model system. The feedback survey was designed to gauge opinions on different components of the system and collect suggestions for improvements. While the survey asked teachers, principals and administrators to respond to a variety of closed-ended questions (analyzed above) concerning their experience with the state model system, there were also three open-ended questions for these educators to consider. The questions were:

1. Please explain your rating of the State Model Evaluation System for Teachers.
2. If you could change the State Model Evaluation System, how would you change it?
3. What additional support would you recommend CDE provide users of the State Model Evaluation System for Teachers?

The purpose of this analysis is to highlight key findings of this open-ended response data in order to provide deeper and more nuanced insight into the experiences of teacher and principals with the state model system.

Summary of key findings

This analysis finds that, overall, teacher and principal responses to open-ended questions concerned: the system itself (documentation, ratings, standards, rubrics, etc.), initial implementation as they experienced it, the potential or ongoing effects of the system in facilitating better teaching practices, evaluation techniques and learning opportunities for students. The following table highlights the dominant¹ themes (in order of prevalence) that emerged from both the teacher and principal responses to the three questions.

¹ At least 10% of the responses must contain the theme in order for it to be considered “dominant” throughout the data set.

Given these prominent themes, this analysis finds that responding teachers were especially concerned about the time it takes to move through the evaluation cycle and the content and clarity of the state model system components themselves (including Bloomboard - an online performance management system that pilot districts are using). Additionally, teachers noted across all three questions the value of differentiating the state model system for teachers in different content areas and grades. While principal comments were generally more positive than those of teachers, they had similar comments and concerns about the system.

QUESTION	TEACHER THEMES (SUBTHEMES) (% OF RESPONSES)	PRINCIPAL THEMES (SUBTHEMES) (% OF RESPONSES)
#1 Please explain your rating of the State Model Evaluation System for Teachers.	Time (time consuming; time taken away from other teaching practices) (21%) Evaluating all teachers on same System (20%) Facilitating teaching and learning practices (14%) Validity (outside of their control; components of the SMES) (14%) Degree of implementation (11%) Reliability (11%)	Facilitating improvement of teacher evaluation (guide teacher growth; feedback conversations) (26%) Time (23%) Validity (14%) Evaluating all teachers on same System (11%) Reliability (11%)
#2 If you could change the State Model Evaluation System, how would you change it?	System components (student assessment; clarity of terms; ratings) (40%) Differentiation for teachers (15%) Size of System (14%) Time (13%) Bloomboard (10%)	System components (30%) Size of System and Time (25%) Bloomboard (23%) Clarity of terms and practices (18%) Differentiation for teachers (14%)
#3 What additional support would you recommend CDE provide users of the State Model Evaluation System for Teachers?	Bloomboard (18%) Professional development (18%) Clarity of components and of implementation (17%) Time (15%)	Professional development (27%) Bloomboard (27%) Examples/exemplars (13%)

Methods of analysis

Among the 813 teachers who completed the survey, 470 (58 percent) answered the first question (“Explain rating of state model system”), 266 (33 percent) answered the second (“Changes to the state model system”) and 333 (41 percent) answered the third (“Support for state model system”). There were 78 principals who completed the survey with 57 (73 percent) answering the first question, 56 (72 percent) answering the second, and 52 (67 percent) answering the third. These responses were then coded thematically in order to determine the most often cited teacher and principal opinions on the System as well as to determine the nuances inherent within these responses. In the first round of coding, several themes emerged from the data.

Notably, while this analysis does not focus on the degree to which responses were positive or negative in nature, approximately 65 percent of the teacher open-ended responses to the first question were determined to be critiquing the system in some way. While this percentage is consistent with the ratings of the state model system in the closed-ended question, *In your opinion, how effective is the State Model Evaluation System for Teachers?*², it is likely the case that teachers withheld more positive comments because of the nature of this survey (it was designed to assess a piloted system seeking improvement). Open-ended questions often provide the space to expand on any concerns the responder may have and thus, they tend to be more skewed toward the negative. This is also likely the case with the principal responses where approximately 49 percent of the principal open-ended responses were critical.

Dominant trends

Six prominent themes emerged from the teacher and principal open ended responses. These were (listed in order of prevalence):

1. Time
2. Evaluating all teachers on same system
3. Facilitating teaching and learning practices
4. Validity or accuracy
5. Degree of implementation
6. Reliability or consistency of ratings

Other notable, but less common, themes that emerged within responses had to do with *Bloomboard* (an online performance management system that pilot districts are using), the *language* used in documentation (such as the rubric and user’s guide) and the *clarity* of the system, particularly in terms of expectations and requirements. In what follows, the dominant themes are defined and quantified³.

Time

Approximately 21 percent of teachers (N=96) who responded cited *time* as a major reason for why they rated the system as they did in the previous, closed-ended question (“How effective is the State Model

² Teachers choosing either “Somewhat effective” or “Not effective” were considered as giving “negative ratings” (while responses of “Effective” and “Very effective” were deemed positive).

³ The majority of responses were categorized under more than one theme.

Evaluation System for Teachers?”). Approximately 14 percent of teachers (N=37) referred to the need to shorten and simplify the SMES in some way. Approximately 13 percent of teachers (N=35) of teachers referred to the need to change the amount of time the state model system took them and others involved (e.g. principals). Within this theme, two subthemes emerged; the system as generally *time consuming* and, specifically, *taking away from other teaching practices*.

Approximately 23 percent of principals (N=13) who responded cited *time* as a major reason for why they rated the system as they did. “It is very time consuming and since most evaluators wear more than one hat, and have multiple tasks to complete, finding the time to complete them is burdensome.” Approximately 25 percent of responding principals (N=14) referred to the need to shorten and simplify the SMES in some way and/or to the need to change the amount of time the state model system took them and others involved.

Evaluating all teachers on the same system

Teachers referred to the ways the system was intended for all teachers but may not be accurately evaluating them based on the particularities of their position. Approximately 20 percent (N=92) of responses contained this theme.

Approximately 11 percent of principals (N=6) who responded referred to the ways the system was intended for all teachers but may not be accurately evaluating them based on the particularities of their position: “*It works well for regular classroom teachers but is difficult for specialists, non-TCAP content, etc.*”

Facilitating teaching and learning practices

Approximately 14 percent of the teachers (N= 64) described ways that the system either promoted or hindered more effective teaching and learning practices: “*Feedback is more specific and allows me to set actionable goals.*”

“*The more effective I am as a teacher, the less this model actually helps me to become a better and more effective teacher. I can see the less effective you are, the better this model is to assist.*”

Approximately 26 percent of principals (N=15) who responded cited the degree to which the system helped to facilitate improved teacher evaluation practices. Within this theme, principals referred positively to how the system will help *guide teacher growth* and, more specifically, the degree to which *feedback conversations* are beneficial for those involved.

Degree of Implementation

Approximately 11 percent of the teachers (N=53) referred to the degree to which the system was implemented with fidelity and/or consistency over the course of the pilot within their district. Teachers stated the need for more clarity from their evaluators around observations (when, how often) and more consistent patterns of feedback.

This theme was not as prominent in the principal responses.

Reliability

Approximately 11 percent of teachers (N=50) of responses referred to issues of reliability with the system or whether the system was appropriately objective and/or could ensure inter-rater reliability: *“I feel like the potential is there, but I believe that it is still very subjective. The rubric is only as good as the conversation that follows.”*

Approximately 11 percent of principals (N=6) of responses referred to issues of reliability with the system or whether the system was appropriately objective and/or could ensure inter-rater reliability: *“The system is only as good as the administrator and teacher using it. The tool doesn't get the result, the individual does.”*

Validity

Approximately 14 percent of teachers (N=67) of responses referred to issues of validity with the system or the extent to which the system was accurately measuring their effectiveness as teachers. Within this theme, teachers were concerned about factors *outside of their control* and with the *components of the state model system* itself: *“There are some parts that I think are especially hard to evaluate because we have no control of what happens outside our walls.”*

Approximately 14 percent of principals (N=8) of responses referred to issues of validity with the system or the extent to which the system was accurately measuring teachers' effectiveness: *“It is a good system and seems to include all elements of a teacher's practice and work with kids and families.”*

Conclusion

This analysis finds that responding teachers and principals were especially concerned about the time it takes to move through the evaluation cycle and the content and clarity of the state model system components themselves. These themes indicate the needs and desires of this sample of teachers and principals as they experienced the state model system for the first time. While there were certainly positive responses throughout the data, these themes tend to reflect areas where growth and change are possible for the state model system. In fact, many of the themes that were identified in the survey were known to CDE staff throughout ongoing work with pilots and other focus groups. As a result, many of the concerns were incorporated into revisions of the state model system rolled out in August 2013 or were scheduled for later rollout. These improvements to the state model system fall under several categories:

- Revisions to the rubric and user guide: These include shortening the rubric, removing redundant professional practices, changing the lowest category from “not-evident” to “basic.”
- Development and roll-out of Elevate Colorado: an inter-rater agreement system to assist evaluators in becoming more consistent.
- Development of [Resource Guide for Deepening the Understanding of Teacher's Professional Practices](#) that defines each practice in the rubric with more specificity and provides examples of the practice.
- Adding additional trainings and webinars to break down the evaluation system components and expectations for more clarity of understanding.

- Development of Excel rubrics for districts wanting a “lower-tech” solution to track data on the evaluations.
- Development of more examples for the measures of student learning side of the equation and video of teachers illustrating high quality professional practice.

Future surveys and analysis will continue to track these themes. This analysis has provided a starting point for improvements already made in the system—and improvements that will still need to be made, over time, as the system evolves.