

Consortium for
Educational
Research and
Evaluation–
North
Carolina

Teacher and Principal Perceptions of the North Carolina Educator Evaluation System

Final Evaluation Report

Authors:

Cassandra R. Davis, Lauren Bangert, Aubrey N. Comperatore, and
Michael Smalenberger

Education Policy Initiative at Carolina, University of North
Carolina at Chapel Hill

September 2015

Consortium for
Educational
Research and
Evaluation–
North
Carolina



Table of Contents

Executive Summary 2

Introduction..... 6

 Brief Overview of the North Carolina Educator Evaluation System..... 6

 Calculation of the Student Achievement Growth Standards 7

 Purpose of this Report..... 8

Data Sources and Methods..... 9

 Methods..... 9

 Sample..... 9

Findings..... 11

 1. How are principals and teachers using measures of student growth (EVAAS) for evaluation and to inform teaching practices? 11

 Evaluation Process and Feedback..... 11

 Teacher Knowledge-Sharing 12

 Data-Driven Instruction 14

 2. What are teachers’ and school leaders’ perceptions of the use of student growth data in the evaluation process? 15

 Knowledge of NCEES 16

 Mixed Perceptions of Using Student Growth Data in the Evaluation 17

 Perceptions of Fairness 17

Summary of Findings..... 23

Conclusion 24

 Recommendations..... 24

References..... 25

Appendix A. Principal Interview Protocols 26

Appendix B. Teacher Interview Protocols..... 31

TEACHER AND PRINCIPAL PERCEPTIONS OF THE NORTH CAROLINA EDUCATOR EVALUATION SYSTEM: FINAL EVALUATION REPORT

Executive Summary

To enhance North Carolina's competitiveness in receiving federal Race to the Top (RttT) funds, the North Carolina State Board of Education agreed to include a measure of student growth for teachers and principals in the existing North Carolina Educator Evaluation System (NCEES). The Consortium for Educational Research and Evaluation–North Carolina (CERE–NC) evaluated the Race to the Top (RttT) initiative to integrate and fully implement the addition of a student growth measure into the NCEES process for teachers. The goal of this evaluation report is to examine teachers' and principals' perceptions of the addition and implementation of the student growth measure to the evaluation process.

Changing the Teacher and Principal Evaluation Processes

Addition of a Student Growth Measure to the Evaluation Process

Prior to RttT, the NCEES, developed with the input of teachers and principals, consisted of five standards for teachers. During RttT, the state adopted a Student Achievement Growth Standard (Standard 6) for teachers that is based on value-added measures of student growth. By including this standard, state, district, and school leaders will be able to assess educator effectiveness using objective measures of student growth in their efforts to improve overall student achievement.

Teachers receive separate ratings on each standard. To be proficient, teachers must receive a rating of proficient or above on the existing standards and a rating of either Meeting Growth Expectations or Exceeding Growth Expectations on the Student Achievement Growth Standard.

In order to receive a rating for the Student Achievement Growth Standard, teachers must generate three consecutive years of student achievement growth scores. School leaders are required to meet with teachers and develop a Performance Development Plan (PDP) for those whose scores indicate inadequate student growth. The United States Department of Education approved an initial implementation timeline in which school year 2012-13 was the first year of data collection; therefore, school year 2014-15 was the first year for which three-year effectiveness scores were available.

Measuring Student Growth

Education Value-Added Assessment System. In 2012, the State Board of Education selected SAS Institute's Education Value-Added Assessment System (EVAAS) to measure student growth. In addition to providing the student growth scores, the online system serves as a single repository for related tools, such as classroom observations, self-assessments, and PDPs. This online system provides educators with access to valuable information across Local Education Agencies (LEAs), regions, and standards. Concurrent with the North Carolina Department of Public Instruction's examination of alternative measures for student growth, CERE–NC conducted an evaluation of the current evaluation system that compared alternative value-added measures of

student growth and found that the EVAAS measure was one of three top-performing approaches.¹

Calculation of the student achievement growth scores. In 2011, the state proposed a plan to calculate teacher effectiveness that was based on a weighted average of individual teachers' value-added scores and school value-added scores using EVAAS estimations.² An analysis of this approach revealed that low-performing teachers in high-performing schools scored higher, while high-performing teachers in low-performing schools received lower scores.³

As a result, in May 2013, the State Board of Education approved an amendment that altered the calculation of teacher effectiveness. A teacher's growth value now is based only on the student growth values for the individual students taught by that teacher. If an education does not have test scores for his or her individual students, the growth value will be based on the data for the entire school. The amendment also determines the effectiveness of principals by including Common Exam data in their evaluation scores.

Methods and Sample

The Evaluation Team conducted 140 structured interviews with principals and teachers during Fall 2013 and Spring 2014. The purpose was to determine the extent to which North Carolina educators are using EVAAS data to inform practices, as well as to collect participants' perspectives on the use of growth data in evaluations.

Evaluation Questions

The findings in this report address the following questions:

1. How are teachers and principals using EVAAS data for evaluation purposes and to inform teaching practices?
2. What are teachers' and school leaders' perceptions of the use of growth data in the evaluation?

Summary of Findings

1. *How are teachers and principals using EVAAS data for evaluation purposes and to inform teaching practices?*
 - a. Teachers and principals appreciated receiving and giving, respectively, feedback on improving instructional strategies, although both agreed that post-evaluation sessions felt rushed and lacked detail about the evaluation scores. Teachers articulated that including information that centered on the meaning of their scores would have been helpful for improving their instructional practices.

¹ Henry et al. (2015); <http://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Evaluation-of-NC-Teacher-Evaluation-9-3-15.pdf>

² The weights chosen were 30% school value-added and 70% individual teacher value-added.

³ Garland, Johnson, & Preston (2013); <http://www.ncpublicschools.org/docs/accountability/common-exams/201303-sbe-update.pdf>

- b. Overall, participants noted that the NCEES helped to increase collaboration between teachers, schools, and districts. Both teachers and principals expressed that the new evaluation system has prompted deeper and more substantive conversations centered on student achievement. In contrast, some participants argued that NCEES has created an unhealthy competitive work environment that has stifled collaboration.
 - c. Generally, teachers said that data pulled from NCEES and other assessments helped them plan for differentiated instruction, highlight areas of strength and weakness within their lessons and instructional practices, and become more reflective educators. Teachers across schools agreed that data from NCEES led to better-informed teaching practices.
2. *What are teachers' and school leaders' perceptions about the use of growth data in the evaluation?*
- a. As indicated in previous reports, both principals and teachers felt unclear about the NCEES measure of student achievement growth. Of the participants who said the tool was useful, the vast majority of principals and teachers alike expressed an interest in receiving additional training that centered on how EVAAS scores were calculated and strategies for improving instructional practices through using student growth data.
 - b. Overall, both teachers and principals shared mixed feelings about the use of student growth data in evaluations. Some teachers agreed that the new process allowed them to fine-tune their instructional practices by addressing weak areas. In contrast, other participants indicated that the data-driven culture reduced flexibility and creativity in the classroom, and prevented teachers from designing instruction based on their own professional knowledge.
 - c. Participants provided recommendations to improve NCEES: 1) account for extenuating circumstances, such as a student's home environment, behavior, and ability; 2) include a section within a standard that evaluates teachers' relationships and interactions with students in- and outside of the classroom; and 3) reduce the weight of the student growth standard. Participants also voiced that the new system created additional stress and pressure on educators.

Recommendations

- *Expand training related to NCEES Standard 6.* Both principals and teachers requested additional information regarding the NCEES and specifically Standard 6, the student achievement growth standard. Principals indicated a desire to have discussions that centered on linking student growth data to instructional practices, while teachers expressed an interest in understanding how administrators calculated Standard 6. Findings reveal that teachers shared similar misconceptions about how the NCEES and Standard 6 were calculated. Some argued that a way to improve the tool would be to include individual growth measures, seemingly revealing their lack of information about the fact that the values for Standard 6 are calculated based on student growth from the previous year. Given these gaps, the Evaluation Team recommends providing more face-to-face and web trainings about how Standard 6 does, in fact, reflect student growth, as well as about how to use EVAAS data to guide instructional improvement. Through these trainings, principals should be better equipped to handle questions about the fairness of NCEES.

- *Continue to seek out teacher input on improving NCEES.* Participants agreed that NCEES could be improved if it accounted for students' extenuating circumstances (such as by including a section evaluating teachers' relationships and interactions with students), and if the weight of the student growth standard were reduced. Participants argued that addressing these components will help create a more accurate assessment of effectiveness.
- *Continue to use feedback loops with teachers.* Occasionally, principals acknowledged an improvement in their teachers' instructional practices and vocally shared these findings with their teachers. Teachers recalled these moments as times that were beneficial to their self-esteem, their profession, and their students. Given the positive response from teachers, the Evaluation Team recommends that principals continue to use the evaluation feedback loop as an opportunity to give positive reinforcement to their staff.
- *Continue to use NCEES as a gateway for teacher collaboration.* NCEES has promoted deeper and more substantive conversations among teachers in Professional Learning Communities (PLCs) Both principals and teachers agreed that these collaborations have improved PLCs. Teachers indicated using information from feedback sessions to assist their peers and jointly improve upon their instructional practices. Some principals noted witnessing a culture shift within their schools that reflected more of a collaborative network.
- *Consider an additional standard that addresses relational aspects of the teaching profession.* Both teachers and principals shared that a missing aspect of NCEES centered on the relationships and bonds teachers form with their students. Participants indicated that the evaluation did not fully recognize everything that a teacher does with her or his students. Research should be used to determine how a relationship component could be documented, tracked, and calculated, potentially including the use of student surveys, as is the practice in some other states.

Introduction

As part of its federal Race to the Top (RttT) proposal, North Carolina committed to integrate and fully implementing the addition of a student growth measure into the North Carolina Educator Evaluation System (NCEES) process for teachers. The Consortium for Educational Research and Evaluation–North Carolina⁴ evaluated several aspects of this integration and implementation over the course of the RttT period.⁵ The purpose of this final report is to determine the extent to which North Carolina educators are using SAS Institute’s Education Value-Added Assessment System (EVAAS) data (the data used in the new NCEES student growth measure) to inform their teaching practices, as well as to collect participants’ perspectives on the use of growth data in evaluations.

Brief Overview of the North Carolina Educator Evaluation System

The NCEES includes both the North Carolina Teacher Evaluation Process (NC TEP) and the Principal Evaluation Process (NC PEP). Prior to the RttT period, principals used the NC TEP to assess teacher performance based on the North Carolina Professional Teaching Standards, guided by an observation rubric with five standards (Table 1, following page) and artifact review.⁶ In February 2012, the State Board of Education, in alignment with North Carolina’s RttT plan, adopted a sixth standard that requires the use of student growth data as a measure of teacher performance.

Similar to the NC TEP, the standards for principals address multiple leadership attributes (strategic, instructional, cultural, human resources, managerial, external development, and micro-political). To inform individual ratings, the NC PEP recommends use of artifacts or evidence (such as data from the Teacher Working Conditions Survey⁷) as well as objective measures (such as teacher turnover) to help principals prioritize and address areas that need additional focus. For principals, measures of student growth were added in 2012 as an eighth standard. Table 1 (following page) includes the title of each standard in the NC TEP and NC PEP.⁸

⁴ CERE-NC is comprised of qualitative and quantitative researchers from the Friday Institute for Educational Innovation at North Carolina State University, the Education Policy Initiative at Carolina at the University of North Carolina-Chapel Hill, and the SERVE Center at the University of North Carolina-Greensboro.

⁵ Previous reports are available at: <http://cerenc.org/rttt-evaluation/teacher-and-principal-effectiveness/>

⁶ See <http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf> for the detailed rubrics used in the NC TEP ratings.

⁷ School-based certified employees are invited to take this survey during even-numbered years. This survey assesses employees’ perceptions of their working conditions. Questions fall within one of the eight categories, time, facilities and sources, community support and involvement, managing student conduct, teacher leadership, school leadership, professional development, and instructional practices and support.

⁸ See <http://www.dpi.state.nc.us/docs/district-humanresources/evaluation/standardsadmin.pdf> for detailed information regarding the NC PEP Standards

Table 1. Standards Addressed in the NC TEP and NC PEP

Standard	NC TEP	NC PEP
1	Teachers demonstrate leadership	Strategic leadership
2	Teachers establish a respectful environment for a diverse population of students	Instructional leadership
3	Teachers know the content they teach	Cultural leadership
4	Teachers facilitate learning for their students	Human resource leadership
5	Teachers reflect on their practice	Managerial leadership
6	Teachers contribute to the academic success of students	External development leadership
7	n/a	Micropolitical leadership
8	n/a	Academic achievement leadership

On both the NC TEP and NC PEP, an educator’s observed level of competence for each standard is rated using a 5-point scale of *Not Demonstrated*, *Developing*, *Proficient*, *Accomplished*, or *Distinguished*. Together, these ratings determine an educator’s overall status, which falls into one of three categories: *Highly Effective*, *Effective*, or *In Need of Improvement*. For example, in order for an educator to earn a rating of *Effective*, she or he must meet expectations for each instructional practice standard, as well as for the student growth standard (Standard 6 for teachers and Standard 8 for principals). An educator who receives a rating of *Highly Effective* must exceed expectations for each standard. All educators rated *In Need of Improvement* must complete a Professional Development Plan (PDP). Only educators who have accrued three consecutive years of data are able to receive an effectiveness status.⁹

Calculation of the Student Achievement Growth Standards

Originally, student achievement growth scores teachers in tested subjects combined teacher- and school-level EVAAS scores. In addition, in 2011, the State Board of Education adopted a policy for calculating teacher effectiveness that also included a measure of school growth (30%) alongside the EVAAS scores (70%). The rationale behind combining components was that teachers and school leaders would be more likely to collaborate if their evaluation scores were tied to school-wide effectiveness.

However, an analysis of this approach revealed that low-performing teachers in high-performing schools scored higher than they should have, while high-performing teachers in low-performing schools received lower scores than they should have.¹⁰ As a result, in May 2013, the State Board of Education approved an amendment that altered the calculation of teacher effectiveness. A teacher’s growth score now is based only on the student growth values for the individual students taught by that teacher. If an educator does not have test scores for his or her individual students, his or her growth value is based on the school EVAAS score alone. The amendment also

⁹ Garland, Johnson, & Preston (2013)

¹⁰ Ibid.

addressed a way to determine the effectiveness of principals by including Common Exam data in their evaluation scores.¹¹

Purpose of this Report

The purpose of this report is to examine educators' perspectives on the new evaluation standards and the effect of these standards on educators' practices and attitudes. To do so, the Evaluation Team conducted qualitative interviews to assess the impact of the evaluation process on educators' attitudes and practices.

The findings in this report address the following research questions:¹²

1. How are teachers and principals using EVAAS data for evaluation purposes and to inform teaching practices?
2. What are teachers' and school leaders' perceptions of the use of growth data in the evaluation?

In addition to these questions, the Evaluation Team also examined participants' views of the professional development resources and information they received, as well as the purpose, process, and effectiveness of their evaluations and the NCEES process as a whole.

¹¹ An earlier report in this series includes more information about the calculation of the student achievement growth standards: Henry et al. (2015); <http://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Evaluation-of-NC-Teacher-Evaluation-9-3-15.pdf>.

¹² Additional information on background objectives and the scope of work for this evaluation are included in previous reports in this series: <http://cerenc.org/rttt-evaluation/teacher-and-principal-effectiveness/>.

Data Sources and Methods

While the evaluation as a whole has used a mixed-methods approach, this report focuses on qualitative data and analysis. The qualitative findings presented in this report are derived from 74 interviews conducted in Fall 2013 and 66 interviews conducted in Spring 2014. Interviews focused on participants' perceptions of the use of student growth data in evaluations, as well as on the types of and processes by which they received evaluative feedback.

Methods

The interview questions pertained to participants' perceptions of the NCEES process, their use of student data, and the types of feedback they gave and received. The Evaluation Team conducted in-person interviews, which were audio recorded for accuracy. In general, the interviews lasted between 45 and 90 minutes. Interviews were then transcribed and uploaded into qualitative analysis software (Atlas.ti Version 7) for analysis.

Initial coding was based on *a priori* codes developed from both the research questions and previous interview analyses. Additional codes were developed through an iterative process and refined based on need and clarity. Two members of the Evaluation Team coded separate interviews with reliability checks interspersed. Inter-rater reliability checks were conducted with both coders and a third member of the Team to ensure consistent coding was employed. Any and all conflicts were discussed and resolved during these checks.

Sample

The Evaluation Team interviewed 52 teachers and 22 of their principals in Fall 2013, and 50 teachers and 16 of their principals in Spring 2014.

The interview sample for Fall 2013 consisted of a purposefully selected group of teachers and principals who were part of a larger sample for the broader evaluation who represented various school contexts and who administered a variety of measures of student learning (e.g., EOGs, EOCs, final exams, etc.).

The selection process for the Spring 2014 sample followed different criteria. For this sample, the Evaluation Team identified participants based on their geographical location (not a part of the criteria for the larger sample from which the Fall 2013 sample was pulled) to ensure a broad representation of the state's educator population. Within these locations, schools were randomly selected across several dimensions: their level (elementary, middle, or high school), their designation as a District and School Transformation (DST) school, and whether they increased student growth by one standard deviation or decreased by one standard deviation over a two-year period. These sampling criteria were imposed as a means of examining any trends in evaluation process and teachers' use of student data across school performance levels. See Table 2 (following page) for a more detailed breakdown of the sample.

Table 2. Interviewed Participants by Year Involved with Research, Role, and School Level

	Teachers		Principals		Total
	Fall 2013	Spring 2014	Fall 2013	Spring 2013	
<i>Elementary</i>	17	15	7	5	44
<i>Middle</i>	18	19	8	6	51
<i>High</i>	16	16	7	5	44
<i>K-8</i>	1	0	0	0	1
Total	52	50	22	16	140

Findings

1. How are principals and teachers using measures of student growth (EVAAS) for evaluation and to inform teaching practices?

In summary, participants regarded the NCEES evaluation process and feedback, as well as the teacher-knowledge-sharing and development of strategies for data-driven instruction that resulted from this process, as methods to help educators inform their teaching practices. This section presents participant descriptions of these processes and practices.

Evaluation Process and Feedback

Overall, participants described the feedback process as a series of interactions between principals and teachers. Principals tended to use the data to establish dialogues with teachers about their performance. Throughout the school year, principals observed what occurred in the classroom, conducted feedback sessions, and then scheduled follow-up discussions with teachers. Participants identified principal walkthroughs, principal observations, and informal and formal conversations between teachers and principals as key components of the evaluation process. In addition to these components, principals indicated using NCEES resources to guide their questions to teachers about their instructional practices. In some instances, principals reported using orientations at the beginning of the year to educate their faculty on the new evaluation system and to address questions about the observation process.

According to both teachers and principals, feedback was focused on teaching practices. Participants stated that each standard was discussed during feedback sessions, along with teacher's assigned rating and evidence for the rating. During feedback sessions, teachers received copies of observers' notes, along with suggestions for improvement. Teachers stated that feedback sessions varied in length and that typically the observers were interested in seeing their lesson plans and how their teaching aligned with objectives. Principals tended to observe less experienced teachers about three or four times a year and seasoned teachers about one or two times.

Overwhelmingly, teachers agreed that their evaluators' feedback was useful because it often included information about how to improve their instruction. Participants identified a shared understanding between evaluator and teacher that feedback on instructional practices could lead to growth in student achievement. Participants noted that administrators provided good and constructive feedback on ways teachers could improve. One teacher acknowledged the importance of having an evaluation conducted by an experienced administrator:

I think it has allowed me to hone in on areas that I need to improve. Especially because it is my administrators who are [more] experienced than me, evaluating me, they can look at what I am doing and say, "This is what you need to work on," and it could be something that I don't necessarily see myself. So, it allows me to be given that feedback and tie it in with [my] PDP [Professional Development Plan] to give me that area to get better.

Participants also indicated that teachers were “buying in” to changing their instructional practices if it meant meeting the needs of students and including more meaningful activities. One teacher stated, “If I get feedback saying that perhaps I could change this or expand that, I am going to give it a try.” Principals also indicated promoting teacher buy-in by targeting teachers’ weakest instructional areas and giving them strategies for focusing on one area instead of on multiple areas. One principal stated:

I believe if teachers get bogged down in trying to do too many things at once, that there could be some issues. So I try to find the thing that they’re the weakest in or the thing that they’re not doing that may help bring them up instructionally-wise and improve student achievement.

Occasionally, principals were able to see and acknowledge improvement in their teachers’ instructional strategies. Teachers recalled these moments as times that were beneficial to their self-esteem, their profession, and their students.

Both principals and teachers shared a wide array of reactions to the timeline and quality of feedback. Principals reported that feedback on observations was comprehensive and immediate. Specifically, principals indicated that feedback occurred within a day or two after observations. In some instances, teachers reported receiving a satisfactory level of feedback on their evaluation, but others indicated a lack of feedback. Although some principals indicated that they fully utilized the feedback loop, others described this process as one that needed the most improvement. Both teachers and principals expressed that their post-evaluation sessions often felt rushed. One teacher stated:

So I would say feedback is probably the area that needs the most improvement just because, like I said, the post-observation meetings, those usually kind of go by the wayside and never really happen, which . . . in my opinion, would be the most important part, because after the observation, that’s when you get that feedback that’s going to be necessary for you to improve. . . . I feel like many times in the past, it’s felt very rushed and it’s also felt just kind of as something to check the boxes, something we have to do, as opposed to something to improve quality.

Teacher Knowledge-Sharing

One important method used by participants to help inform teaching practices was sharing information with their peers about how to improve student achievement. Overwhelmingly, teachers articulated that due to the changes in NCEES, they were more likely to collaborate with teachers within their schools and across the district. In addition, both teachers and principals noted a positive change in their schools’ cultures. In a few instances, however, teachers said that they were not able to collaborate with their peers, usually because of limited resources and limited time, and some participants even contended that NCEES promoted unhealthy competition between teachers and thus diminished collaborative efforts.

Most participants said that the addition of student growth measures to the NCEES helped to increase collaboration between teachers. Principals noted that teachers became more willing to collaborate with their peers and less likely to remain isolated as a result of the incorporation of

student data. One principal stated, “Teachers aren’t going in their rooms and shutting their doors anymore. They’re coming out; they’re talking to one another. So I think that is a good thing.” Another teacher noted the importance of sharing knowledge about instructional practices in order to meet the needs of all students in her classroom. The teacher stated:

I collaborate so much more with my teammates. We are all trying to just be on one accord and just do what we can do to exchange ideas and exchange different strategies and different teach[ing] methods, just so we can try to make sure that all the learners in our classroom are being taught and being covered.

Participants also stated that the NCEES has prompted deeper and more substantive conversations between teachers during Professional Learning Community (PLC) meetings. Both principals and teachers indicated that PLC meetings in the past mainly focused on addressing administration needs. With the help of data pulled from the NCEES, participants were able to have thorough discussions about meeting the needs of students and teachers. One teacher recognized the change in the structure of PLC meetings: “So, [PLCs] are actually having conversations now, so it is different, it is different.”

In addition to increased collaborations among teachers, both teachers and principals agreed that NCEES has helped improve knowledge-sharing between schools. In some instances, schools compared NCEES results across schools, which resulted in teachers expressing an interest in visiting and seeing what more successful schools were doing in their classrooms. Participants agreed that results from NCEES opened the door for more cross-district PLC collaboration on deconstructing the standards and making them applicable to all educators.

Due to the increase in collaboration, both teachers and principals articulated a change in the school culture that resulted in a communicative environment for educators. Principals said that the new collaborative culture was a beneficial shift for teachers because it forced all educators to discuss strategies for specific students. One principal described a PLC meeting in which he noticed all teachers engaged in meeting the needs of one student. The principal stated:

The [teachers] would talk about what they were doing in each classroom and whether they were being successful, why they weren’t being successful. If they were being successful here but they’re not over here, [and ask questions like,] “What are we doing? What can we do?”

Principals noted that teachers consistently disseminated information about students’ needs. Participants indicated that they distributed student information more readily in places outside of traditional PLC meetings. Both principals and teachers agreed that these collaborations helped to inform PLC conversations. In some instances, teachers stated that they collaborated with other teachers in their grades every week, and once a month, participants shared best practices to ensure alignment of lessons. One teacher stated:

My grade level works real closely together, so we meet a lot. I mean, we are constantly sharing ideas, we plan together. . . . We met with second grade at the last one to do some alignment discussion—you know, just to see that we are all on the same page. So, we do a lot of communicating and collaborating.

Although most teachers agreed that the NCEES allowed for more opportunities to collaborate, some teachers reported not having the chance to collaborate due to busy schedules and limited planning time. Teachers and principals alike expressed that teachers did not collaborate as much as they needed to. In a few instances, teachers affirmed that they did not have a PLC for an entire year. In their opinion, the collaboration between teachers and schools also has dropped off. One teacher stated:

The only time [we] really . . . have time for collaboration is on our own at lunch—whoever has lunch with us—or if we have a common planning period, which doesn't happen very often. They do try to design some professional development opportunities in our district where teachers—say, high school math teachers—are grouped together, or middle school math teachers are grouped together, but we have had an awful lot of change in our calendar this year. . . . So, I feel like we have really missed on some opportunities there this year.

Some principals and teachers noted that the effects of Standard 6 on personnel and pay decisions has led to unhealthy competitions amongst teachers and lowered collaboration. Several principals shared that the recent decision to align growth data to merit pay created a negative and reclusive school environment. One principal argued that the new aggressive climate has caused teachers to keep their strategies to themselves, which ultimately diminished collaboration between teachers. One teacher stated:

[The system has] made us competitive, and so if I know only twenty-five percent of the teachers are going to get a contract based on Standard 6, there's a real contradiction that I'm supposed to be in the PLC and working together.

Another teacher argued that NCEES created an atmosphere that was not beneficial to the teaching profession. The teacher stated, “It creates a climate of competition and I personally don't feel that that is an effective climate for the profession that we are in.”

Data-Driven Instruction

Another key method participants used to inform teaching practices focused on using EVAAS data to help drive instruction. Overall, teachers noted that data pulled from the NCEES and other assessments helped them plan for differentiated instruction, highlight areas of strength and weakness within their curricula, and become more reflective educators. Teachers across schools talked about how students in their classrooms were at varying learning levels, but that they could and did use data to create learning clusters, group tutoring sessions, and individual progress check-ins with students. One teacher spoke about using data to determine which of her students needed additional help: “I use that information to plan my guided reading groups and then plan the students that will go into which group and then to plan for each specific group.” Another teacher spoke about using the data consistently to determine how to best plan for classroom instruction. The teacher stated, “I use [data] pretty much on a daily basis. [I use] the data I get out of the classroom to know how I'm going to teach the next day.”

Teachers also mentioned using data to help drive instructional changes within their schools. One teacher stated, “As a school, we use it to guide our intervention block. We use that data to

determine the skills of the students while we are working on them.” In one instance, a teacher described how the entire grade level split its students up into learning groups based on data compiled from assessments. The teacher recalled each teacher supporting her or his group of students and using data to track individual and group progress.

In addition to using data to plan for differentiated instruction, teachers recalled using data from NCEES to highlight strengths and weaknesses among certain student groups. Through the use of student data, teachers were able to determine where students were unclear of content, and then change their instructional practices based on their findings. Teachers across schools insisted that the data from NCEES led to better-informed teaching practices. One teacher stated that low testing scores from the gifted/academically intelligent [AI] students forced teachers throughout the school to seek a new way to support this group of students. The teacher stated, “Like, one year as a school, we saw that we were all weak with our AI kids. [They] weren’t rising like they were [AI], so we kind of put a push towards that and just doing higher-level thinking skills with them especially.”

Lastly, participants reported that teachers were more reflective practitioners as a result of increasing their use of and dependency on student data. In some cases, principals felt that their teachers had become more “diagnostic” and “prescriptive” in their analysis of student deficiencies and application of interventions as a result of collaboration. In other instances, teachers were more likely to make data-driven decisions. These decisions helped the teachers identify where they could improve in their instruction and lesson-planning. Teachers and principals alike agreed that if teachers improved, then their students would improve as well. One teacher expressed being dependent on data from the NCEES in order to directly impact instruction and student progress:

Everything for me is based on data. . . . I grow from my data. . . . It makes me very conscious of the fact that, “Oh, this area, I have got to target this area and make sure that I have some plan and make sure I am covering it.” So, if there is any area or any particular point that wasn’t quite high enough, I am like, “Okay, I have got to [hit] that really hard. . . .” So, it does affect my teacher practice, and it affects the students because they are going to get more from me than they did before.

Another teacher stated, “[Using data] has helped me [be] more reflective, both on the lessons I do and my interactions and reactions with kids. It has helped me to become more proactive with noticing kids’ misunderstanding or misconceptions.”

2. What are teachers’ and school leaders’ perceptions of the use of student growth data in the evaluation process?

Interview data revealed several key themes related to participants’ perceptions of the use of student growth data in the evaluation process. In summary, participants expressed having limited knowledge of the NCEES and shared mixed sentiments about the use of student growth data in teacher evaluations. In addition, teachers and principals both suggested strategies to improve the instrument’s fairness and also expressed concerns about its impact on educator morale.

Knowledge of NCEES

When asked to provide thoughts on their knowledge of NCEES, participants were more likely to explain Standards 1 through 5 rather than Standard 6. In some instances, teachers had a solid understanding of what was being looked for in Standards 1 through 5, but were unsure about how they were being evaluated on Standard 6. One teacher stated, “I feel like I have a general understanding of what’s being looked at, but . . . [the standards] are not explained very clearly.” Additionally, principals shared mixed feelings about their teachers’ knowledge of the evaluation process. In some instances, principals believed that their teachers lacked knowledge about NCEES, while in other cases, principals assumed that their teachers were fairly knowledgeable about the system. One principal stated:

[My teachers] understand their Standards. They know what the expectations are. The one they’re really familiar with, I’d say, [is] Standard 6 because they know that’s where the growth is, and they’re going to receive that from the state. But Standards 1 through 5 they’re knowledgeable in that.”

Principals also indicated that they believed that NCEES would provide greater benefit to their teachers if they had more clarity on the overall evaluation tool. One principal stated, “I think the benefits would be greater if we had a better way and a consistent way to help people understand what the numbers even mean.” Some principals indicated that teachers had the most questions about how teacher effectiveness and the formula for calculating a year’s growth were determined. One principal stated:

I think understanding how they came up with whether or not the teacher was considered effective or not effective, and how they came up with the formula for whether or not a child has made the year’s worth of growth and things like that—that’s what my teachers have had the most questions about.

Overall, teachers indicated an interest in receiving more professional development and evaluative feedback, especially regarding Standard 6. One teacher stated, “I feel like [the evaluation] just needs to be condensed somewhere and I think we need more training on Standard 6 because I honestly don’t know what really to do with it.”

Another teacher reported:

When we get the evaluations back, it’s checks in boxes. So, knowing which [standards] to pay attention to if there’s one that we really . . . need to focus on, I think that [could] be better explained. . . . Right now, I think a lot of teachers, and myself included sometimes, see it as a hoop to jump through instead of something meaningful for self-reflection.

These concerns may be related to an admitted lack of understanding or knowledge regarding the use of Standard 6 data in their evaluations. One teacher stated:

Standard 6 is . . . one of those things that brings about a lot of heartburn and headaches and stress headaches to a lot of different people, because nobody really knows what’s going on with Standard 6.

Perceptions of Using Student Growth Data in the Evaluation

Another theme centered on participants' mixed sentiments about the use of student growth data in teacher evaluations.

As indicated earlier, some teachers perceived the new system—including Standard 6—to be a positive tool that would improve their teaching practices. Many agreed that the adjustment to the new evaluation system was difficult, but they maintained that data pulled from the system ultimately would help them support their students. Some teachers expressed a hopeful outlook that the use of student growth data would allow teachers to be held accountable for their instructional practices and challenge them to become more effective teachers. Specifically, some teachers reported that having access to student growth data would be beneficial to overall teacher morale. Teachers described being motivated by being able to see their students grow, and by being able to know where students were in their learning. One teacher stated, “The times where you feel like you are getting pulled in so many different directions, it is good to see that the kids are still learning and growing, and looking at data helps me to know that.” Another teacher commented, “I think it will encourage teachers to continue to strive to do better. . . . Seeing those numbers does encourage you.” One principal expressed support of Standards 6, and said that the assessment focused on student growth and not proficiency:

We are looking to see how students grow, and for each year of education of schooling students ought to grow, and I don't think that's an expectation that is unreasonable, and so I do think it should be an integral part of teacher performance, their evaluation.

Although many participants expressed positive views of the NCEES, in contrast some principals and teachers disagreed with the use of Standard 6, and viewed it as an unnecessary and biased assessment. One principal felt it was not appropriate to use student test scores for evaluating a teacher. The school leader deemed all principals to be responsible for determining if a teacher were effective or not, and not the job of Standard 6. For this principal, Standard 6 represented a movement that disenfranchised principals from being able to lead their schools:

So, my responsibility is to recruit and hire the most effective individual in the classroom, and if they're not effective, that's my job to get in there and take care of it. It should not be the job of the student's test score to say whether or not that teacher is effective.

Perceptions of Fairness

Participants also discussed the overall fairness of NCEES, and included suggestions for ways to improve results. Participants shared three ways to make the instrument fairer: 1) account for extenuating circumstances, such as a student's home environment, behavior, and ability; 2) include a section in one of the standards that evaluates teachers' relationships and interactions with students inside and outside of the classroom; and 3) reduce the weight of the student growth standard to less than that of the remaining standards.¹³ Finally, participants noted how NCEES

¹³ Some participants stated that individual student growth data should be included in EVAAS scores, a suggestion that is noteworthy largely because, as indicated earlier, it suggests that some teachers do not fully understand how EVAAS scores are calculated.

has impacted their morale as an educator. Each of these reflections is explored in more detail in this section.

First, participants argued that student growth data does not take into account the life experiences of students that may impact their ability to perform in the classroom. Some of the issues mentioned by teachers included students' home environments, socio-economic status, family situations, behavior, attendance, and disabilities. Both teachers and principals emphasized that a student's home life, especially for economically disadvantaged students, may affect how well she or he can function at school. One teacher stated:

You don't know if these kids had to go home and had to babysit all night or what the deal is, because especially in a poor county like this, you don't know what kind of home life these kids have, and it impacts what happens here [at school].

Participants said that test scores do not account for extenuating circumstances, such as whether a student is a poor tester. One teacher argued that the evaluation fails to capture whether a student is having a bad day and performs poorly as a result. In this case, the teacher is left with a low evaluation score based on a student's *testing* ability, instead of based on a student's *academic* ability. One teacher stated, "Well, you never know [whether] that kid [is] going to test well that day—they may or they may not. So, it may be good, and they do well, so there you go, that looks good for you; [or] it could be a bad day, so they don't test well." In spite of the fact that teachers' EVAAS scores are based on growth rather than on levels of test scores, some teachers insisted that students who lag behind at the beginning of the year prevent them from having an accurate assessment of teacher performance:

A lot of the students that we get are so far behind the standards that we're holding them to. For instance, if I get a student and one of the standards is to maybe get them through multiplication, and the student is still struggling with addition, then it's not really a fair measure if I'm being judged or critiqued on how well I got the student through this standard, [because] the student came to me at a very low level and I brought them up from that level.

In addition to addressing outside factors such as home life, testing ability, and achievement level, teachers also argued that one's ability to manage one's classroom greatly impacts evaluation scores. Teachers that dealt with behavior issues in the classroom may still be effective teachers; however, having to give these behavior issues more attention could have negatively impacted their teaching practices. As one participant put it, "You have some teachers that have behavior problems in the classroom. She can't help the situation if [there are] behavior problems in the classroom, but she is still [having to teach], too." Another teacher explained how student behavior may affect a teacher's overall effectiveness: "Some days you *can't* teach because you are dealing with student behavior. So, how will that affect my effectiveness?"

Many teachers suggested that it was unjust to base a teacher's performance on the academic achievement of their students given these extenuating factors that are outside of their control. One teacher stated:

I'm torn, because in a perfect world if you're doing your job, [students] are going to grow, but in a non-perfect world—which is what we live in—there are so many aspects in their life outside of school that impedes learning, that prevents them from growing. . . . The only one that's responsible and that will be evaluated on that is the teacher.

Nearly all participants argued that it was possible for teachers' evaluations to be good based on Standards 1 through 5 but still have low Standard 6 score because students are dealing with outside factors that cause them to test poorly. One principal stated that his school had 90% of students qualifying for free and reduced lunch and that the school personnel struggled to meet the basic needs of their students. The principal argued that other priorities often trump the standards:

We have to expose students to things that they would not normally be exposed to, and with that challenge comes a balancing issue between how much time is spent on the standards versus how much time is spent bringing students up to the level [at which] they are ready to be able to demonstrate proficiency with the standards.

In addition to accounting for extenuating circumstances, the second approach participants mentioned as a way to improve NCEES was by including additional components to assess during the evaluation, such as outside classroom activities and teacher-student relationships. Principals and teachers alike acknowledged that the evaluations did not fully recognize everything a teacher did for her or his students, and that including these components would help paint a clearer picture of the teachers' role and impact.

Participants frequently stated that the NCEES does not take into account a teacher's efforts outside of the classroom to help students, such as providing extra tutoring, purchasing classroom supplies, feeding students, and working on the weekends. Both principals and teachers expressed that the evaluation should include these missing components in the teacher ratings. One teacher stated:

I think there's a lot of behind-the-scenes [work] that teachers invest in that never gets evaluated with evaluations, and I realize they have to be done in some manner but I don't know that forty-five minutes four times a year truly covers what a teacher does.

Another teacher emphasized that including these missing pieces would help define the overall snapshot of the role of a teacher: "When you see people out here on Saturdays and you see us out here until 5:00, 5:30 tutoring kids and taking kids home, that should play a part in [how] we do at the end of the year [on evaluations]."

Both principals and teachers agreed that the NCEES fails to account for teacher-student relationships inside and outside of the classroom. Some participants argued that the evaluation tool should include a component that evaluates these relationships. One principal recalled recent research on student engagement that emphasized the importance of teacher-student relationships:

I think there needs to be more additional components that deals with the stuff that teachers do besides just what they do in the classroom. Because again, if you look at a lot of the recent literature going out there, one of the ways to improve student engagement is for the kids to know that you care about them, not just about their chemistry scores, but

that they are on the baseball team or that their sister is getting married on Saturday. So, there needs to be, I think, more of a focus on that appropriate student-teacher relationship. What are they doing outside of the class to work [for] their kids? Because it is amazing, my teacher of the year also has the most clubs and the largest clubs.

Participants argued that before an educator can teach students, she or he must develop a relationship with her or his students. If a teacher has a good relationship with a student, it will motivate the student to work harder. One teacher stated:

If that student can trust the teacher, if the student knows that the teacher is there to help him or her, I think they are more willing to try harder and perform better, and you cannot see that. You cannot see how hard some of them are willing to stay after school and put in the extra time and succeed and keep trying and keep trying and keep trying.

Some participants expressed doubt, however, that, although a teacher's rapport with students is vital, this phenomenon could be measured. One principal stated, "You can't measure the rapport that the teacher has with their students and the relationship they develop with the students in the classroom and how that helps motivate a student."

The third way participants suggested for improving the fairness of the NCEES was shifting the weight of the standards. Teachers had mixed opinions regarding the use of Standard 6 in their evaluations. Some teachers advocated an equal weighting of all of the standards: "I think [the standards] should carry equal weight, with the understanding that if a teacher is lacking in one area, it doesn't mean that they are a bad teacher." Others, however, preferred that Standards 1 through 5 be weighted more heavily than Standard 6. As one teacher contended, doing so would ensure that Standard 6 would not be solely used to assess a teacher's effectiveness:

I don't ultimately think they should be equally weighted. I think that Standard 6 should be weighted, but not as high. . . . I don't think a test should really . . . be the end-all be-all if you are an effective teacher. So, I think that Standards 1 through 5 should be more weighed than Standard 6.

Principals strongly believed that effective teachers who instruct low-achieving students might achieve high ratings on Standards 1 through 5, but still end up with a low Standard 6 rating, while an equally effective teacher who instructs high-achieving students may achieve comparable ratings on Standards 1 through 5 and a high rating on Standard 6. Therefore, the discrepancy in Standard 6 ratings may not stem from the effectiveness of the teacher, but from the cohort of students.

The final approach participants suggested for improving the fairness of NCEES was to change the way the state estimates student growth, though their suggestions in most cases primarily illustrated the need for clarifying how EVAAS scores are calculated. For example, several teachers shared a misconception that EVAAS scores do not take into consideration how much the student has grown, instead only measuring whether or not the student is performing at a predetermined level of proficiency. For example, one teacher asserted that her definition of progress differed from the state's perception:

[Progress] is the student who comes to you and they can barely do certain things that you are asking them to do and then at the end of time span, they are doing amazing things just from where they started, [even though maybe not where] the state or other people say they should be.

Some teachers perceived that it was harder to show growth for students who were already high-achievers, and that data from these students might not accurately reflect the abilities of the teacher. Other teachers stated that the evaluation should emphasize the growth of the individual student so that the performances of teachers who have primarily low-achieving students are not compared to teachers who have primarily high-achieving students: “[NCEES] should emphasize the growth of the individual student, so that if a student comes in low and then grows at the end of the year, than the results will show that the teacher’s instruction impacted the student.”

While these suggestions reflect for the most part how EVAAS scores already are calculated, the concerns these teachers raise may be salient when considering teachers who instruct students with disabilities. Students with disabilities may express what they have learned in different ways than students without disabilities, and hence different measures of student growth beyond written tests would be more appropriate for these students:

I still feel that the measures that we’re using—especially for special education children who may actually gain a lot of knowledge and growth even within the curriculum but can’t express their learning in ways that are typical of children who do not have disabilities—I feel that the way we measure is very unfair to the children, and using those measures for growth is not fair because . . . their disability may disguise what their learning is.

In addition to participants’ strategies for improving the fairness of NCEES, teachers and principals both agreed that the new system created additional stress and pressure on educators. Teachers reported several concerns related to the use of student growth data in their evaluation, including increased stress, added pressure, and lower levels of self-efficacy. One teacher expressed, “I just think it’s a long process, very time-consuming, not only for the administrator but [also] for the teachers. . . . It’s just overwhelming to me.”

Teachers described how their morale, and view of themselves, may be negatively affected when their teacher ratings are lower or higher than expected. Some teachers reported experiencing discouragement as a result of scores that did not match their expectations. One teacher stated:

Some of our teachers, when they looked up their scores, it changed their whole morale. Even the ones that did very well and the ones that didn’t do so very well. Every teacher is a teacher for a reason, and it takes a different breed of a person to be a teacher. So, I think we really take it to heart, thinking that we either made a kid’s year or ruined the kids’ year. I think that hits us hard, and I think that can impact, it really can impact.

Additionally, teachers expressed concerns related to increased stress and an awareness of “considerably more pressure” to promote student growth, both for themselves and their students. Many teachers found it difficult not to “take it personally” when it comes to their student growth

data. For many teachers, this feeling was a result not only of wanting to be successful for their evaluation but also of wanting to do right by their students. One teacher commented:

They feel considerably more pressure because they're aware of the data. . . . [I]t's not just about self-preservation. It's about they don't want to do their children wrong. They want to serve their children well, and if the data shows they're not growing the way they should, they really take it personally, take it to heart.

Teachers indicated that the evaluation places an emphasis on ensuring that the curriculum aligns with required tests, which reduces flexibility and creativity in the classroom. Teachers reported concerns that they are not doing a good job since they are primarily focusing on student data, which has required them to “waste time on test-taking strategies” instead of concentrating on effectively teaching the material to their students. One teacher stated:

I could have [students] take practice test after practice test, and that's not what I want to do. I want them to have fun and research, but in terms of student growth data . . . I feel like once again it's taking teaching out of it and turning this into a business.

In addition, principals argued that NCEES has created urgency for teachers to teach to the test, and that this effort is forcing students to complete more work in a shorter period of time. One principal argued that the greatest impact of NCEES on teachers was in pacing. The administrator noted that teachers felt the pressure to cover all elements of the curriculum:

I've heard a lot of teachers talking about [how] one standard may have five components, and they feel like they've got to . . . cover it all. You have to, and I think it's a greater sense of urgency and maybe in some classrooms a certain panic that they're just not going to get it done in time from a pacing perspective.

Although principals noted more teacher stress, in some instances principals argued that the heightened stress has not impacted their teachers' instructional strategies. One principal stated:

The only thing that I have seen is that there is more stress, that [teachers] feel that they [have] got to [finish] this part, too. [However,] I don't think that they have changed the way they have taught, that they are working any harder; they give their best every day and I don't think that that standard has changed that.”

Summary of Findings

1. *How are teachers and principals using EVAAS data for evaluation purposes and to inform teaching practices?*
 - a. Teachers and principals appreciated receiving and giving, respectively, feedback on improving instructional strategies, although both agreed that post-evaluation sessions felt rushed and lacked detail about the evaluation scores. Teachers articulated that including information that centered on the meaning of their scores would have been helpful for improving their instructional practices.
 - b. Overall, participants noted that the NCEES helped to increase collaboration between teachers, schools, and districts. Both teachers and principals expressed that the new evaluation system has prompted deeper and more substantive conversations centered on student achievement. In contrast, some participants argued that NCEES has created an unhealthy competitive work environment that has stifled collaboration.
 - c. Generally, teachers said that data pulled from NCEES and other assessments helped them plan for differentiated instruction, highlight areas of strength and weakness within their lessons and instructional practices, and become more reflective educators. Teachers across schools agreed that data from NCEES led to better-informed teaching practices.

2. *What are teachers' and school leaders' perceptions of the use of growth data in the evaluation?*
 - a. As indicated in previous reports, both principals and teachers felt unclear about the NCEES measure of student achievement growth. Of the participants who said the tool was useful, the vast majority of principals and teachers alike expressed an interest in receiving additional training that centered on how EVAAS scores were calculated and strategies for improving instructional practices through using student growth data.
 - b. Overall, both teachers and principals shared mixed feelings about the use of student growth data in evaluations. Some teachers agreed that the new process allowed them to fine-tune their instructional practices by addressing weak areas. In contrast, other participants indicated that the data-driven culture reduced flexibility and creativity in the classroom, and prevented teachers from designing instruction based on their own professional knowledge.
 - c. Participants provided recommendations to improve NCEES: 1) account for extenuating circumstances, such as a student's home environment, behavior, and ability; 2) include a section within a standard that evaluates teachers' relationships and interactions with students in- and outside of the classroom; and 3) reduce the weight of the student growth standard. Participants also voiced that the new system created additional stress and pressure on educators.

Conclusion

Recommendations

- *Expand training related to NCEES Standard 6.* Both principals and teachers requested additional information regarding the NCEES and specifically Standard 6, the student achievement growth standard. Principals indicated a desire to have discussions that centered on linking student growth data to instructional practices, while teachers expressed an interest in understanding how administrators calculated Standard 6. Findings reveal that teachers shared similar misconceptions about how the NCEES and Standard 6 were calculated. Some argued that a way to improve the tool would be to include individual growth measures, seemingly revealing their lack of information about the fact that the values for Standard 6 are calculated based on student growth from the previous year. Given these gaps, the Evaluation Team recommends providing more face-to-face and web trainings about how Standard 6 does, in fact, reflect student growth, as well as about how to use EVAAS data to guide instructional improvement. Through these trainings, principals should be better equipped to handle questions about the fairness of NCEES.
- *Continue to seek out teacher input on improving NCEES.* Participants agreed that NCEES could be improved if it accounted for students' extenuating circumstances (such as by including a section evaluating teachers' relationships and interactions with students), and if the weight of the student growth standard were reduced. Participants argued that addressing these components will help create a more accurate assessment of effectiveness.
- *Continue to use feedback loops with teachers.* Occasionally, principals acknowledged an improvement in their teachers' instructional practices and vocally shared these findings with their teachers. Teachers recalled these moments as times that were beneficial to their self-esteem, their profession, and their students. Given the positive response from teachers, the Evaluation Team recommends that principals continue to use the evaluation feedback loop as an opportunity to give positive reinforcement to their staff.
- *Continue to use NCEES as a gateway for teacher collaboration.* NCEES has promoted deeper and more substantive conversations among teachers in Professional Learning Communities (PLCs) Both principals and teachers agreed that these collaborations have improved PLCs. Teachers indicated using information from feedback sessions to assist their peers and jointly improve upon their instructional practices. Some principals noted witnessing a culture shift within their schools that reflected more of a collaborative network.
- *Consider an additional standard that addresses relational aspects of the teaching profession.* Both teachers and principals shared that a missing aspect of NCEES centered on the relationships and bonds teachers form with their students. Participants indicated that the evaluation did not fully recognize everything that a teacher does with her or his students. Research should be used to determine how a relationship component could be documented, tracked, and calculated, potentially including the use of student surveys, as is the practice in some other states.

References

- Garland, R., Johnson, L., and Preston, J. (2013). *North Carolina educator effectiveness policies and update*. Retrieved from <http://www.ncpublicschools.org/docs/accountability/common-exams/201303-sbe-update.pdf>.
- Garland, R. and Quick, A. (2013). *MSL/Common Exam spring 2013 update*. Retrieved from <http://www.ncpublicschools.org/docs/effectiveness-model/resources/2013/02-memo.pdf>.
- Henry, G.T., Guthrie, J.E., and Davis, C.R. (2015). An evaluation of the North Carolina educator evaluation system and the student achievement growth standards. Retrieved from <http://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Evaluation-of-NC-Teacher-Evaluation-9-3-15.pdf>.
- Measures of Student Learning: NC's Common Exams. (2013) Retrieved from <http://www.ncpublicschools.org/docs/accountability/common-exams/implementation-guide.pdf>.
- North Carolina Department of Public Instruction. (2013). North Carolina standards for school executives. Retrieved from <http://www.dpi.state.nc.us/docs/district-humanresources/evaluation/standardsadmin.pdf>.
- North Carolina Department of Public Instruction. (2012). North Carolina teacher evaluation process. Retrieved from <http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf>.
- North Carolina Race to the Top Proposal (2010) Retrieved from <http://www.ncpublicschools.org/rttt/state/>.

Appendix A. Principal Interview Protocols

Survey Questions Used from Interview Protocol (Principal) Fall 2013

(11) Please share any updates about the types of information you have received on the addition of student growth data to the *Teacher and/or Principal Evaluation Process*?

→ *PROBES*:

- Describe what information you have received through online resources, such as modules, webinars, and evaluation dashboards.
- Describe what you still feel you need/want to know.

(12) Describe your experience with EVAAS student growth data, if any, in this current school year.

→ *PROBES*:

- To what extent has it been used to provide feedback to and/or evaluation of teaching?
- To what extent has it been used to inform instruction?

(13) Tell us about the process of implementation of Common Exams in your LEA last year (2012-13).

→ *PROBES*:

- How were they administered (ex. Online, face-to-face, combination)?
- Who scored the exams?
- What were challenges in implementation, if any?
- What lessons did you learn from last year's implementation?

(14) Describe your experience with Common Exam data to date.

→ *PROBES*:

- To what extent has it been used to provide feedback to and/or evaluation of teaching?
- To what extent has it been used to inform instruction?

(15) What are your current feelings about the use of student growth data in evaluations?

→ *PROBES*:

- What potential challenges do you see for the addition of these new standards to the Teacher and Principal Evaluation Process?
- What do you see as the potential benefits?

(16) To what extent do you feel the addition of student growth data in teacher evaluations measure teacher effectiveness?

→ *PROBES*:

- How do these ratings on Standards 1-5 compared with Standard 6 help to differentiate between each teacher's effectiveness?
- What are your thoughts on the relationship between Standards 1-5 and Standard 6? For instance, can a teacher achieve a high score on Standard 6 without similar scores in the other five standards?

(17) If you had to identify your most effective teachers, are there any additional criteria you would use aside from the NCEES?

→ *PROBES*:

- Are there other indicators of teacher effectiveness NOT captured in the current evaluation system?
- What do you think teachers do that is important that may NOT be considered in the teacher evaluation?

(18) What changes, if any, have you observed in teachers' instructional practices as a result of the addition of student growth data in the evaluation process?

→ *PROBES*:

- What changes, if any, have you see in educator collaboration as a result of the addition of student growth data in the evaluation process?

(19) Describe any areas that you would appreciate further professional development related to the new teacher evaluation standard, EVAAS, and the addition of other MSLs, such as the Common Exams.

(20) What other questions, concerns, or comments would you like to share with us at this time?

Survey Questions Used from Interview Protocol (Principal) Spring 2014

(1) To start, please tell us what subject(s) you have taught this year. Please give us the grade level(s) and the names of the class(es).

(2) Walk us through your teacher evaluation process. Please describe all the ways you are evaluated from the beginning of the school year through the end of the school year.

→ *PROBES*:

- Who evaluates you?
- By what methods? (observation, review of curriculum/student work, etc.)
- How often and at what points during the school year do formal evaluations occur?
- How often and at what points during the school year do informal evaluations occur (such as classroom walk-throughs).

(3) Please describe the *process for receiving feedback* related to your evaluation.

→*PROBES*:

- Who provides you feedback related to your evaluations?
- When and how often do you receive feedback (e.g. when during the year and at what times, such as before or after an observation, other times)?
- How is this feedback provided? (written, conference, combination, etc.)

(4) Please describe the *content of the feedback* you receive related to your evaluation.

→*PROBES*:

- To what extent is feedback connected to your teaching practices?
- To what extent is feedback connected to your students' learning and growth?
- To what extent do you receive feedback on the progress you've made since your last evaluation?

(5) To what extent is your evaluation feedback directly connected to the NC Teaching Standards?

Reminder: The standards are: (1) Demonstrating leadership; (2) Establishing a respectful environment for a diverse population of students; (3) Content knowledge; (4) Facilitating learning for students; (5) Reflecting on practice; and (6) Contributing to the academic success of students.

→*PROBES*:

- In which of these standards do you tend to receive the MOST feedback?
- In which of these standards do you tend to receive the LEAST feedback?

(6) What kinds of resources and/or professional development support do *teachers* at your school receive related to their evaluation?

→*PROBES*:

- Do teachers at your school receive any coaching?
- Do teachers have [regular] progress checks at your school?
- Do teachers collaborate to improve their practices, such as through instructional rounds or Professional Learning Communities (PLCs)?

(7) Please describe any resources and/or professional development that *you* have been provided in relation to your evaluation.

→*PROBES*:

- Who has provided the resources, at what time points, and how often?
- To what extent are these resources directly related to feedback you received around areas in need of improvement?
- Describe any areas that you would appreciate further professional development related to the teacher evaluation process. (Prompt: NCEES, the addition of student growth measures, EVAAS, Final Exams, etc.)

(8) To what extent do you feel you understand the criteria that are being used for each of your ratings on Standards 1-5?

→*PROBES*:

- To what extent do you feel the criteria for Standards 1-5 can readily be observed within your teaching practices?

(9) To what extent do you understand how the 6th Standard / student growth rating is calculated?

→*PROBES*:

- How long has your school been using EVAAS data to measure student growth?
- If your 6th standard rating is based on an EVAAS score, please describe how your principal (or administrator) supports your understanding of EVAAS data (i.e. school/district PD, online training, electronic communications etc.)
- How has your knowledge/use of EVAAS data changed over time?

(10) There are currently multiple assessments that can be used to calculate a teacher's 6th Standard rating: EOC/EOGs, Final Exams, and Analysis of Student Work. What are your perceptions of each one as a measure of student growth and as a standard for teacher effectiveness?

→*PROBES*:

- To what extent do you think these different assessments authentically measure your impact on student growth in your subject area?
- To what extent do you feel these different assessments provide similar and/or unique information about student growth?
- To what extent do you feel these different assessments provide similar and/or unique information about *teacher effectiveness*?

(11) *ASW Teachers*: Describe your experience thus far in the Analysis of Student Work pilot.

→*PROBES*:

- When did you join the pilot?
- Describe the process for collecting and submitting student evidence in the pilot.
- Describe the review process for Analysis of Student Work.

(12) *ASW Teachers*: What are your perceptions about the Analysis of Student Work as a way to calculate your Standard 6 rating?

→*PROBES*:

- To what extent do you feel Analysis of Student Work authentically measures your impact on student growth in your subject area?
- To what extent do you feel Analysis of Student Work provides consistent information about student growth in your subject area?
- What changes, if any, would you make to the ASW process?

(13) *ASW Teachers*: What are your perceptions about the training experience with the Analysis of Student Work?

- What aspects of the training experience have been most effective?
- What aspects of the training experience would you change, if any?

(14) What are your thoughts about the relationship between Standards 1-5 and Standard 6 as measures of teacher effectiveness?

→*PROBES*:

- To what extent do you feel a teacher can have high ratings on Standards 1-5 but low student growth scores for Standard 6 and vice versa?
- To what extent do you feel these standards should be equally weighted?

(15) To what extent do you feel the current evaluation system represents your effectiveness as a teacher?

→*PROBES*:

- What things do you think teachers do that is important that may NOT be considered in the teacher evaluation?
- To what extent has the teacher evaluation process help you become a better teacher? (e.g. targeted improvement within teaching standards, helping students perform better on tests, etc.)

(16) What have you learned about yourself and/or your students based on your 6th *Standard* / student growth data?

→*PROBES*:

- To what extent does 6th Standard data inform your teaching practices?
- Tell us about your use of other data related to student growth or progress, excluding data used to calculate a 6th Standard rating (EOC/EOGs, Final Exams, and ASWs). For example, utilizing student benchmark assessments.
- To what extent does other data inform your teaching practices?

(17) What changes, if any, would you make to the current teacher evaluation system?

(18) What other questions, concerns, or comments would you like to share with us at this time?

Appendix B. Teacher Interview Protocols

Survey Questions Used from Interview Protocol (Teacher)

Fall 2013

(10) Please share any updates about the types of information you have received on the addition of student growth data to the Teacher Evaluation Process?

→ *PROBES*:

- Describe what information you have received through online resources, such as modules, webinars, and evaluation dashboards.
- Describe what information has been most useful to you.
- Describe what you still feel you need/want to know.

(11) (If they have accessed EVAAS – if not skip to 12)... Describe your experience with EVAAS student growth data, if any, in this current school year.

→ *PROBES*:

- To what extent has it been used to provide feedback to and/or evaluation of your teaching?
- To what extent has it been used to inform instruction?

(12) (If they have accessed Common Exams – if not skip to 13)... Describe your experience with Common Exams to date.

→ *PROBES*:

- To what extent has it been used to provide feedback to and/or evaluation of your teaching?
- To what extent has it been used to inform instruction?

(13) Describe your experience with additional student growth data, excluding Common Exams, EOCs/EOGs.

→ *PROBES*

- What other student growth data do you use?
- To what extent has it been used to inform instruction?

(14) What are your current feelings about the use of student growth data (i.e. MSLs) in teacher evaluations?

→*PROBES*:

- What potential challenges do you see for the addition of this new standard to the Teacher Evaluation Process?
- What do you see as the potential benefits?
- Do you get feedback related to student growth data in a timely fashion (e.g. in time to make adjustments to your instruction)?

(15) To what extent do you feel the addition of student growth data in teacher evaluations impact teacher effectiveness?

→*PROBES*:

- Are there other indicators NOT captured in the current evaluation system that also impact effectiveness? (for example, an indicator could be improved student behavior)
- What do you think teachers do that is important that may NOT be considered in the teacher evaluation?

(16) What changes, if any, have you made in your instruction as a result of the addition of student growth data in the evaluation process?

(17) What changes, if any, have you seen in educator collaboration as a result of the addition of student growth data in the evaluation process?

(18) Describe any areas that you would appreciate further professional development related to the new teacher evaluation standard, EVAAS, and the addition of other MSLs, such as the Common Exams.

(19) What other questions, concerns, or comments would you like to share with us at this time?

Survey Questions Used from Interview Protocol (Teacher) Spring 2014

(1) To start, please tell us what subject(s) you have taught this year. Please give us the grade level(s) and the names of the class(es).

(2) Walk us through your teacher evaluation process. Please describe all the ways you are evaluated from the beginning of the school year through the end of the school year.

→*PROBES*:

- Who evaluates you?
- By what methods? (observation, review of curriculum/student work, etc.)
- How often and at what points during the school year do formal evaluations occur?
- How often and at what points during the school year do informal evaluations occur (such as classroom walk-throughs).

(3) Please describe the *process for receiving feedback* related to your evaluation.

→*PROBES*:

- Who provides you feedback related to your evaluations?
- When and how often do you receive feedback (e.g. when during the year and at what times, such as before or after an observation, other times)?
- How is this feedback provided? (written, conference, combination, etc.)

(4) Please describe the *content of the feedback* you receive related to your evaluation.

→*PROBES*:

- To what extent is feedback connected to your teaching practices?
- To what extent is feedback connected to your students' learning and growth?
- To what extent do you receive feedback on the progress you've made since your last evaluation?

(5) To what extent is your evaluation feedback directly connected to the NC Teaching Standards?

Reminder: The standards are: (1) Demonstrating leadership; (2) Establishing a respectful environment for a diverse population of students; (3) Content knowledge; (4) Facilitating learning for students; (5) Reflecting on practice; and (6) Contributing to the academic success of students.

→*PROBES*:

- In which of these standards do you tend to receive the MOST feedback?
- In which of these standards do you tend to receive the LEAST feedback?

(6) What kinds of resources and/or professional development support do *teachers* at your school receive related to their evaluation?

→*PROBES*:

- Do teachers at your school receive any coaching?
- Do teachers have [regular] progress checks at your school?
- Do teachers collaborate to improve their practices, such as through instructional rounds or Professional Learning Communities (PLCs)?

(7) Please describe any resources and/or professional development that *you* have been provided in relation to your evaluation.

→*PROBES*:

- Who has provided the resources, at what time points, and how often?
- To what extent are these resources directly related to feedback you received around areas in need of improvement?
- Describe any areas that you would appreciate further professional development related to the teacher evaluation process. (Prompt: NCEES, the addition of student growth measures, EVAAS, Final Exams, etc.)

(8) To what extent do you feel you understand the criteria that are being used for each of your ratings on Standards 1-5?

→*PROBES*:

- To what extent do you feel the criteria for Standards 1-5 can readily be observed within your teaching practices?

(9) To what extent do you understand how the 6th Standard / student growth rating is calculated?

→*PROBES*:

- How long has your school been using EVAAS data to measure student growth?
- If your 6th standard rating is based on an EVAAS score, please describe how your principal (or administrator) supports your understanding of EVAAS data (i.e. school/district PD, online training, electronic communications etc.)
- How has your knowledge/use of EVAAS data changed over time?

(10) There are currently multiple assessments that can be used to calculate a teacher's 6th Standard rating: EOC/EOGs, Final Exams, and Analysis of Student Work. What are your perceptions of each one as a measure of student growth and as a standard for teacher effectiveness?

→*PROBES*:

- To what extent do you think these different assessments authentically measure your impact on student growth in your subject area?
- To what extent do you feel these different assessments provide similar and/or unique information about student growth?
- To what extent do you feel these different assessments provide similar and/or unique information about teacher effectiveness?

(11) *ASW Teachers*: Describe your experience thus far in the Analysis of Student Work pilot.

→*PROBES*:

- When did you join the pilot?
- Describe the process for collecting and submitting student evidence in the pilot.
- Describe the review process for Analysis of Student Work.

(12) *ASW Teachers*: What are your perceptions about the Analysis of Student Work as a way to calculate your Standard 6 rating?

→*PROBES*:

- To what extent do you feel Analysis of Student Work authentically measures your impact on student growth in your subject area?
- To what extent do you feel Analysis of Student Work provides consistent information about student growth in your subject area?
- What changes, if any, would you make to the ASW process?

(13) *ASW Teachers*: What are your perceptions about the training experience with the Analysis of Student Work?

- What aspects of the training experience have been most effective?
- What aspects of the training experience would you change, if any?

(14) What are your thoughts about the relationship between Standards 1-5 and Standard 6 as measures of teacher effectiveness?

→*PROBES*:

- To what extent do you feel a teacher can have high ratings on Standards 1-5 but low student growth scores for Standard 6 and vice versa?
- To what extent do you feel these standards should be equally weighted?

(15) To what extent do you feel the current evaluation system represents your effectiveness as a teacher?

→*PROBES*:

- What things do you think teachers do that is important that may NOT be considered in the teacher evaluation?
- To what extent has the teacher evaluation process help you become a better teacher? (e.g. targeted improvement within teaching standards, helping students perform better on tests, etc.)

(16) What have you learned about yourself and/or your students based on your 6th *Standard* / student growth data?

→*PROBES*:

- To what extent does 6th Standard data inform your teaching practices?
- Tell us about your use of other data related to student growth or progress, excluding data used to calculate a 6th Standard rating (EOC/EOGs, Final Exams, and ASWs). For example, utilizing student benchmark assessments.
- To what extent does other data inform your teaching practices?

(17) What changes, if any, would you make to the current teacher evaluation system?

(18) What other questions, concerns, or comments would you like to share with us at this time?

Contact Information:
Please direct all inquiries to Cassandra Davis
cnrichar@email.unc.edu

© 2015 Consortium for Educational Research and Evaluation–North Carolina



● Carolina Institute
for Public Policy



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

