

Consortium for  
Educational  
Research and  
Evaluation–  
North  
Carolina

*Executive Summary of*

# An Evaluation of the North Carolina Educator Evaluation System and the Student Achievement Growth Standard

2010-11 through 2013-14

Gary T. Henry and J. Edward Guthrie  
Vanderbilt University

September 2015

Consortium for  
Educational  
Research and  
Evaluation–  
North  
Carolina



Carolina Institute for Public Policy  
THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL



# AN EVALUATION OF THE NORTH CAROLINA EDUCATOR EVALUATION SYSTEM AND THE STUDENT ACHIEVEMENT GROWTH STANDARD: 2010-11 THROUGH 2013-14

## Executive Summary

In 2011, as a part of the State Board of Education's implementation of North Carolina's Race to the Top (RttT) initiative, a sixth standard—a measure of student growth, the Educational Value-Added Assessment System (EVAAS)—was added to the existing five standards for evaluating teachers. The purpose of this report is to describe the outcomes of teacher evaluations that have occurred since the sixth standard was added and trends in those outcomes through 2013-14.

## *Evaluation Questions and Key Findings*

1. *How many teachers needed improvement, according to the NCEES? To what extent did principals rate teachers below proficient (Not Demonstrated or Developing) or did teachers not meet expected growth according to their EVAAS scores?*

One of the most important purposes of teachers' evaluations is to identify teachers who need improvement so leadership can intervene in ways that help ensure that students have access to high-quality teaching. Between 2010-11 and 2013-14, for all teachers with both individual EVAAS growth scores and principals' ratings, 18.4 percent were found to need improvement, but this percentage varied between initially licensed teachers (21.4 percent) and fully licensed teachers (13.5 percent).

## *Key Findings*

- Over 80 percent of teachers in each of the past two years who were assigned to a "Needs Improvement" category (*Not Demonstrated* or *Developing*) were assigned by their EVAAS score alone.
  - Approximately 5 percent of beginning teachers who received ratings on all five standards received one or more ratings below *Proficient* from their principals.
  - The percentages of teachers who received at least one rating by their principals below *Proficient* was 7.7 percent for teachers who did not meet expected growth according to EVAAS, 3.2 percent for teachers who met expected growth, and 1.2 percent for teachers who exceeded expected growth.
2. *Did principals provide teachers with information on their strengths and weaknesses by making distinctions in performance between the standards?*

Another important purpose for the NCEES evaluations was to provide teachers, especially initially licensed teachers who are, on average, less effective and more likely to turnover than more experienced teachers, with clear information about their strengths and weaknesses. Value-added scores, including but not limited to EVAAS scores, provide objective measures of the outcomes of teachers' instructional practices but, unfortunately, do not provide information about which practices are strengths and weaknesses for individual teachers.

*Key Findings*

- Principals rated teachers either *Proficient* or *Accomplished* 90 percent of the time, which provided limited information on individual teachers' strengths and weaknesses.
  - Principals' ratings have not varied over time, indicating little refinement in using NCEES ratings to provide teachers with feedback on strengths and weaknesses.
  - Principals rate teachers globally rather than providing meaningful distinctions on teachers' performance on each standard.
3. *Were the EVAAS scores for teachers' contributions to student achievement growth related to principals' ratings of teachers' performance, especially the ratings for facilitating student learning?*

To enhance credibility of NCEES with teachers and to provide consistent information, the agreement between the principal ratings that teachers received and their student growth ratings (EVAAS) should be related. The objective measure of their students' achievement growth could reasonably be expected to relate to principals' ratings of teachers on the NCEES standards, especially the "teachers facilitate learning for their students" standard, which most closely has a bearing on student growth.

*Key Findings*

- Principals' ratings of teachers' overall performance and of their ability to facilitate student learning were loosely correlated with teachers' EVAAS scores. On average, principals' ratings of their teachers were not influenced by the state's official measure of the teachers' contributions to the achievement growth of their students.
  - The correlations between principal ratings and EVAAS scores varied across individual course subjects. Correlations with EVAAS were higher with science and substantially lower with reading/English/language arts, the mClass, and career and technical assessments.
4. *What instructional practices or teacher behaviors predicted gains or improvements in teachers' EVAAS scores?*

Teachers need actionable evidence about effective instructional practices that can improve their EVAAS scores. In interviews, most teachers reported receiving immediate and constructive feedback regarding their evaluations; however, some reported a lack of feedback from their administrators. Overall, teachers expressed an interest in receiving a higher quality and greater quantity of feedback.

In order to support principals and teachers in the development of more effective teachers, the Evaluation Team examined measures of teachers' instructional practices and behaviors that relate to improvements in their EVAAS scores—that is, related to higher EVAAS scores than predicted based on their prior EVAAS scores alone. Practices found to be effective and ineffective at improving EVAAS scores, as well as practices associated with teachers' capacity to improve their EVAAS scores over time, are listed in Table ES1 (following page).

Table ES1. Measures Associated with Impacting Teachers' Value-Added Scores More than Predicted by Prior Value-Added Scores

<i>Effective Practices</i>	<i>Practices Associated with Teachers' Capacity to Improve</i>
Facilitating Student Learning	Reflection on Practice
Collaborative Environment	Collaborative Environment
Higher-Order Instruction	Higher-Order Instruction
Classroom Management	Classroom Management
Positive Climate	

*Ineffective Practices*

- Busy Work
- Student-Led Environment

### Key Findings

- Both teachers and principals want more information about how teachers can improve their practices, especially in ways that increase their value-added scores.
  - Using measures from principals' ratings, student surveys, and classroom observations we found several measures associated with higher value-added scores than were predicted by teachers' previous value-added scores alone.
  - These measures can be (1) directly incorporated into the feedback provided by principals to teachers, and (2) periodically tested for their relationship to value-added if student surveys and/or observation protocols for classrooms are conducted in the future.
5. *What were teachers' views about their evaluations and related topics during the period in which the NCEES evaluation with six standards has been implemented?*

When the student achievement growth measure was added to the original five NCEES standards, two concerns surfaced. The first concern was that the overall fairness of the evaluations would be eroded—that the protocols designed to promote teacher development would be undermined. Second, concerns were raised that the high-stakes evaluations would inhibit teachers from supporting one another and working together to improve student learning.

From 2011-12 to 2013-14, the favorability of teachers' views of their evaluations has declined significantly. Overall, teachers' rating of the evaluation process as measured by survey items declined from approximately 5.2, which indicated slight agreement with the items, to about 4.8, which moved them toward neither agreement nor disagreement with the items. Contrary to some expectations, teachers engaged in knowledge-sharing more in 2013-14 than in earlier years. Overall, their knowledge-sharing activities increased from once or twice per week to almost daily.

### Conclusions

Along with approximately two-thirds of other states, North Carolina adopted value-added scores for individual teachers as an additional, sixth standard to supplement the pre-existing five-standard NCEES teacher evaluation system.

- Because almost all identifications of a need for improvement were based on the value-added score, value-added effectively acted alone to determine teachers' evaluation status, rendering the judgments of principals on all other aspects of teaching and teachers' performance much less important.
- Consideration should be given to systematically adding other direct measures of teaching performance into the NCEES in addition to the EVAAS scores that are currently included. In addition to providing more evidence for teachers who also have EVAAS scores, teachers who do not have EVAAS scores could then have direct measures of their individual performance incorporated into their evaluations.
- The second main conclusion from this evaluation is that teachers are being rated globally, and neither the ratings nor evaluation feedback are providing them with enough actionable information for them to improve.
- The NCEES process seems to have been accepted by teachers and principals. On the whole, they feel that the system is fair. Also, teachers' survey responses indicated that the implementation of the NCEES process has not produced negative side effects, such as decreasing teachers' willingness to share information. In fact, the opposite has occurred—teachers have engaged in more knowledge-sharing across the first four years of expanded NCEES implementation.

The main issue with the NCEES appears to be that the current system includes only one systematic data source—EVAAS—and, while EVAAS is an important objective measure of teachers' effectiveness, the inclusion of additional systematic measures may point out the strengths and weaknesses of individual teachers, increase the accuracy of identifying those who need improvement, increase the favorability of teachers' attitudes toward the evaluation system, and provide direct information about practices that can be used for improvement for all teachers.

**Contact Information:**  
Please direct all inquiries to Gary T. Henry  
[Gary.henry@vanderbilt.edu](mailto:Gary.henry@vanderbilt.edu)

© 2015 Consortium for Educational Research and Evaluation–North Carolina



● Carolina Institute  
for Public Policy



THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL ●

