Student Surveys and Teacher Effectiveness

By Joe Maugeri, Julia Burrus Pierson, Vincent Reitano, and Qi Wang Xing

Project 9.2 / August 2015

EXECUTIVE SUMMARY

In 2012, North Carolina piloted a student survey system, through which 25% of LEAs surveyed their students to gather student-level perception data regarding the effectiveness of teachers. This study provides strong evidence of the reliability of student surveys, which is to say that student surveys are consistent in their results. The findings of the predictive models outlined in this report can be translated into a straightforward framework for improving teacher effectiveness. It is recommended that student surveys be used statewide as part of the teacher evaluation and improvement process.

POLICY QUESTIONS

This report explores the relationship between teacher value-added scores and teacher evaluation scores measured via students surveys. Specifically, this report seeks to answer three research questions:

1. Are student surveys internally reliable?
2. Is there a relationship between teacher evaluation scores measured by student surveys and teacher effectiveness as measured by EVAAS?
3. If the answers to the first two questions are yes, are there particular teacher qualities as measured by student surveys that have a stronger relationship with teacher-value added scores? Are there particular qualities that do not?

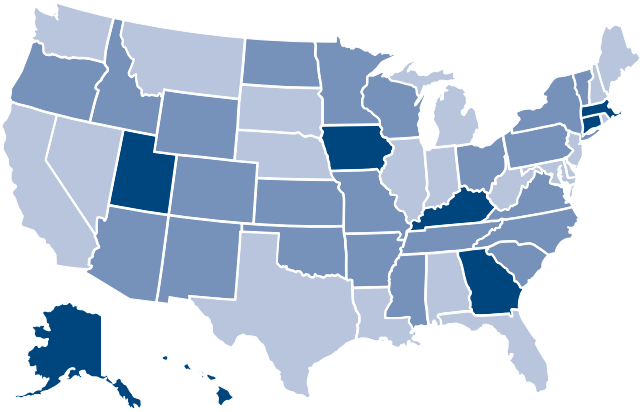
INTRODUCTION

The practice of evaluating teachers has presented unique challenges to statewide education systems. States across the country have redesigned their teacher evaluation systems in response to the Race to the Top (RttT) and Elementary and Secondary Education Act (ESEA) waiver policies. These state plans often present two common components: a value-added model (VAM) that measures teachers' contributions to student growth, and ratings of instructional quality using administrator observations.

STUDENT SURVEYS

Traditionally, individual school administrators, such as principals and assistant principals, evaluate teachers using qualitative measures. Student perception surveys, more often conducted in higher education to rate professors, have gained more attention in recent years in K-12 education, as educational researchers and policymakers began to realize the need for alternative and more comprehensive measures of teacher effectiveness. An increasing number of states and districts are considering utilizing student surveys as instruments in their teacher evaluation systems (Hanover, 2013). As of 2013, 8 states require student surveys in teacher evaluation system, 23 states neither require nor prohibit student surveys, and 20 states explicitly do not permit student surveys (National Council on Teacher Quality, 2013), as displayed in Figure 1. Currently, the North Carolina General Statues neither require nor prohibit the use of student surveys as part of the teacher evaluation process.

Figure 1. State Policies (Student Surveys)



STUDENT SURVEY POLICIES

Require student survey: AK, CA, CT, GA, HI, IA, IL, KY, MA, UT

Allow student surveys: AR, AZ, CO, DC, ID, KS, MN, MO, MS, NC, ND, NM, NY, OH, OK, OR, PA, SC, TN, VA, VT, WI, WY

Explicitly do no permit student surveys: AL, DE, FL, IN, LA, MD, ME, MI, MT, NE, NH, NJ, NV, RI, SD, TX, WA, WV

Sources: National Council on Teacher Quality, Education Commissioner of States

NORTH CAROLINA: TRIPOD AND THE 7CS

In 2012, the North Carolina Department of Public Instruction partnered with Cambridge Education to launch a pilot student survey program in many of North Carolina's school districts. During a four-week window in spring 2012, over 125,000 students in grades K-12 completed an online or paper survey on their learning environment. Students completed either a short or long version of the Tripod Survey, which was designed by Dr. Ron Ferguson from Harvard University. The Bill and Melinda Gates Foundation used a version of the Tripod Survey in its Measures of Effective Teaching Project.

Tripod survey measures are research-validated and based on the conceptual model of the "7Cs." The 7Cs of effective teaching involve a set of best practices that peer-reviewed research has linked to student engagement (effort and behavior) and achievement (gains on standardized tests). The 7Cs framework of effective teaching describes the extent to which teachers:

- CARE: show concern and commitment
- CONFER: invite ideas and promote discussion
- CAPTIVATE: inspire curiosity and interest
- CLARIFY: cultivate understanding and overcome confusion
- CONSOLIDATE: integrate ideas and summarize key points
- CHALLENGE: Press for rigor and persistence
- CONTROL: Sustain order, respect, and focus

EVAAS AND VALUE-ADDED DATA IN NORTH CAROLINA

North Carolina uses the Education Value Added Assessment System (EVAAS) from the SAS Institute as its value-added metric to measure student growth. EVAAS applies a combination of statistical models to assess LEA, school, and individual teacher effectiveness based on student growth. Beginning in 2011-12, EVAAS data became part of the North Carolina Educator Evaluation System for teachers and school administrators, although EVAAS data was being supplied to LEAs prior to this point. Beginning with the 2012-13 school year, the State started reporting on EVAAS data as part of the school accountability model.

METHODOLOGY

DATA

To answer the research questions, unreleased data from the North Carolina Department of Public Instruction student survey pilot project in 2012 is used. It was conducted across a random sample of 25% of all schools, spanning K-2, 3-5, and 6-12. For this analysis, K-2 surveys are excluded because, as of 2012 (the year in which the survey was administered) teacher EVAAS scores are not available for these grades due to an insufficient number of years of data.

Survey Data

For this study, two separate sets of Tripod survey data are used, separated by elementary and secondary grade levels. The elementary sample included 54,643 students in grades 3-5 and the secondary sample included 39,628 students in grades 6-12. For both the elementary and secondary datasets, the short and long versions of the Tripod are combined and any questions that were not on both versions were excluded. In total, there were 42 questions on the compiled elementary survey, 27 of them capturing the 7Cs. There were 65 questions on the compiled secondary survey, 35 of them capturing the 7Cs.

Value-Added Data

To facilitate this research, SAS provided EVAAS data for all North Carolina public school teachers with value-added scores. The data included each teacher's school, LEA, grade level, content area, and value-added score from school year 2011-2012. The data also included the school-level demographic information for percentage of economically disadvantaged students and the percentage of minority students.

Final Dataset

The final datasets for this project were created by merging the student survey data with the teacher value-added data via teacher UIDs, so that all students and their survey answers matched the teachers they surveyed. All teachers who were not assigned an EVAAS score were dropped, as only instructors teaching courses with an EOG or EOC assessment could receive an EVAAS value-added score in 2012. Also dropped were all teachers whose students were not participants in the pilot survey. The final elementary dataset included 1,519 teachers evaluated by 33,762 students and the final secondary dataset included 1,209 teachers evaluated by 32,266 students. The final dataset provides a representative sampling of North Carolina teachers. A new variable for each 7C question was added in order to create a

teacher composite score by question (the composite variable is simply the average score that each individual teacher received on each 7C question). Other school variables, class, and teacher characteristics were included as covariates, such as percentage of economically disadvantaged students, percentage of minority students, class size, teacher absences in days, and a categorical variable to control for subject.

METHODS

With one year of data from the school year 2011-2012, a factor analysis and multinomial logistic regression were conducted to examine the relationship between student survey data and teacher performance. Because of the large number of survey questions, the factor analysis is used mostly for data reduction purposes. A factor analysis is used to condense the survey questions into key, unobserved constructs (or themes) representing a subset of the 7Cs of teaching as defined by TRIPOD: challenge, control, care, confer, captivate, clarify, and consolidate. The factor analysis reveals differences between elementary and secondary data. In particular, with an alpha cutoff of .5 (i.e., any question with a loading score less than .5 would be dropped), the results of the factor analysis yield a simple factor structure with three factors (or clusters of common questions) for elementary surveys (grade 3 to 5): academic support, classroom management, and academic press ***(see Table 1 for elementary factor loadings)***. However, for secondary (grade 6 to 12), the dimensional reduction reveals two factors: academic support and classroom management ***(see Table 2 for secondary factor loadings)***. Academic support includes many of the questions under academic press[1].

However, for secondary survey questions, the dimensional reduction reveals two factors, academic support and classroom management. Academic support includes many of the questions under academic press. This implies that there may be redundancy in the survey for secondary students, which is evidenced by a Cronbach's Alpha above .95 for secondary data. Therefore, the 7Cs are reliable, but may suffer from redundancy.

Because each teacher was rated by multiple students, each teacher's average rating per question was calculated. All analysis is conducted at the aggregated teacher level, in which students are aggregated to each respective teacher at the class level. With the aggregated data and controls for teacher absence, as well as school and class characteristics, teachers' EVAAS scores are regressed on the unobserved constructs for elementary data and secondary data respectively. The outputs from the ordinary least square regressions yield consistently an R-square at the .2 level, indicating that student survey may explain 20% of the variation in teacher's EVAAS score, which suggests that, while exceptional in measuring the value-added of teachers relative to test scores, EVAAS may not capture everything that is conventionally classified as excellent teaching.

Multinomial logistic regression is used when the outcome is a categorical variable, in this case, teacher's EVAAS status (1=In Need of Improvement, 2=Meeting Expectation, 3=Exceeding Expectation), to assess the effect of each construct in predicting teacher effectiveness measured by VAM through EVAAS. This represents a first attempt at modeling how much variation in VAM can be accounted for by student surveys.

$$EVAAS_status = \alpha + \beta_1 * FACTORS + \beta_2 * School + \beta_3 * Class + \beta_4 * Absence + \epsilon$$

EVAAS_status is a 3 point scale of teacher's EVAAS score. Teachers with EVAAS scores below -2 are considered "in need of improvement". Teachers with EVAAS scores in the range of -2 and 2 are considered "meeting expectations". Teachers with EVAAS scores above 2 are considered "exceeding expectations". *FACTORS* is a vector of the key, unobserved constructs from factor analysis. The three constructs from elementary student surveys represent A, B, and C. The two constructs from secondary student surveys represent A, and B. *School* is a vector consistent of percentage of economically disadvantaged students, and percentage of minority students. *Class* is a vector that controls for classroom level covariates, such as class size, grade, and subject. *Absence* is the number of days a teacher was absent during the school year. ϵ is the error term.

The focus is then placed on the marginal effects on the probabilities of a change in the regressors for a given teacher's EVAAS status (Cameron & Trivedi, 2005). Average marginal effects are computed to permit interpretation of coefficients as the percentage change in a given teacher effectiveness status due to one unit change in the control variable. The outputs show that student surveys hold a stronger relationship with EVAAS at the extremes of the distribution of value-added teacher effectiveness. Student ratings of teachers are also found to be more divergent for math and science teachers than it is for English teachers, cautioning careful interpretation of the student ratings by subject.

FINDINGS

DESCRIPTIVE STATISTICS

The descriptive statistics ***(see Table 3 for descriptive statistics)*** generated from the final datasets display trends in teacher EVAAS scores, overall effectiveness measure, student survey data, as well as school and class characteristics (Table 2). Secondary teachers from the sample, on average, have higher EVAAS scores. However, teacher effectiveness, a categorical variable which captures a teacher's overall effectiveness across Standards 1-6 (the six standards through which teachers are evaluated, with standards 1-5 representing domains evaluated through principal observations and standard 6 representing a teacher's EVAAS score), shows that secondary teachers, on average, have slightly lower ratings. This

[1] *23 out of 29 elementary survey questions are reduced into three key, unobserved constructs, and 31 out of 35 secondary survey questions into two key, unobserved constructs. The secondary survey questions initially loaded onto three key factors, however Factor 3 contained only one variable which loaded more strongly onto Factor 1. This variable was dropped and only two factors were retained.*

may indicate elementary teachers' overall rating from Standards 1 to 5 are slightly higher than secondary teachers. The descriptive statistics also display a greater dispersion of EVAAS scores among secondary teachers, which, when compared to the secondary survey results may suggest a 1) greater variance in how secondary school students perceive their teachers relative to elementary students, and 2) the ability of the survey to better capture the perspectives of older and more mature students who can form a more comprehensive evaluation of their teachers.

REGRESSION ESTIMATES

The multinomial logistic regression results show that student surveys are most predictive of performance for teachers in the bottom and top performance groupings ***(see Table 4)***. Marginal effects are estimated at the mean of each variable. Looking at the constructs, classroom management and academic press have moderate effects in magnitude and are statistically significant in predicting teachers' EVAAS rating for elementary teachers. For example, for each one unit increase in classroom management (e.g., increasing a teacher's average survey score in classroom management by one point on the Likert scale used by the survey), the probability of a teacher being rated as "does not meet expected growth" decrease by 4.8%, whereas one unit increase in academic press increase the probability of a teacher being rated as "exceeds expected growth" by 5.4%. The estimates are similar for secondary teachers where academic support and classroom management have moderate effects in magnitude and are statistically significant. This indicates that constructs from student surveys across instructional levels have a positive association with teachers' EVAAS ratings.

The marginal effect estimates on subject area yield interesting results. The subject area English language arts (ELA) is omitted from the estimation and serves as a base model. Therefore, the interpretation of estimates on math and science are relative to that of ELA, which is interpreted as a 0. Across instructional levels, math and science teachers are more likely to be in the top and bottom groups compared to their ELA counterparts, whereas they are more likely to be rated as "meet expected growth". The magnitude of effects are much larger for secondary teachers, which are also statistically significant. For example, a math teacher in secondary schools are 13.9% more likely to be rated as "does not meet expected growth" and 16.5% more likely to be rated as "exceeds expected growth" than ELA teachers. The same holds true for science teachers as well. The effects of school characteristics, such as percentage of economically disadvantaged students and percentage of minority students in secondary schools, while significant, are so small in magnitude that little to no inference can be drawn from them.

DISCUSSION

Overall, the effects of the constructs from student surveys are consistent across instructional levels regarding the effect size and significance level, and there is a positive relationship between student survey and teacher EVAAS status. This suggests that improvement in performance on student surveys would reduce the probability of those teachers scoring in the bottom grouping of EVAAS. Simply put, this means that improvement in the domains measured through student surveys should yield a positive, corresponding impact on teacher EVAAS performance. In addition, the fact that student surveys hold a stronger relationship with EVAAS at the extremes of the distribution of value-added teacher effectiveness may imply that student surveys maintain the most relevance for teachers who are at the bottom and top of growth as measured by EVAAS.

A second major finding pertains to area of improvement of the current student surveys. Larger effects in magnitude with statistical significance are found in secondary science and math. This may imply that TRIPOD is less sensitive to capturing the complexity of subject-specific differences relative to ELA classes. Specifically, students are rating teachers of technical subjects (math and science) with a more divisive approach than ELA, a probable function of the stricter math and science learning standards, which tend to be narrower in focus than ELA standards. Students may be penalizing teachers of these subjects not due to their teaching effectiveness, but rather the difficulty assessing competency in these respective content areas. This is also supported by the lack of significance on elementary school estimates, which is likely attributed to the fact that teachers often cover multiple subject areas at an elementary level relative to middle or high school grades.

POLICY RECOMMENDATIONS

USE STUDENT SURVEYS FOR ALL TEACHERS

Because student surveys have been demonstrated to produce reliable results and these results can be used to predict teacher EVAAS scores, it is recommended that the use of student surveys be expanded to all teachers in North Carolina and that the results of these surveys are used as part of the teacher evaluation process. Student surveys provide reliable data that can be used to help improve teacher effectiveness. The implementation and management of a statewide student survey system will undoubtedly cost money, but it is believed that statewide use of student surveys is a reasonable investment with the goal of providing valuable feedback to teachers.

TARGET STUDENT SURVEY RESULTS ON IMPROVING LOW EVAAS TEACHERS

The results of this study are most pronounced among teachers either "not meeting expected growth" or exceeding expected growth" as measured by EVAAS. The marginal effects of each factor generated for teachers falling in these categories tend to suggest that a change in their survey results would have the greatest impact on their EVAAS scores. Student surveys should be used as a diagnostic tool for all teachers, but the results of this study suggest that the use of student survey results

among the lowest performing teachers relative to EVAAS will help educators to realize the most substantial gains. Because teachers categorized by EVAAS as "exceeding expected growth" are already demonstrably proficient teachers, it is recommended that student survey results be targeted on improving the craft of low EVAAS teachers, which, in turn, could reduce the probability of these teachers being classified as "not meeting expected growth" in subsequent years.

MODIFY SURVEY QUESTIONS TO REFLECT CONTENT-SPECIFIC INSTRUCTION/PEDAGOGY

Student survey results appear to be the most polarizing among math and science teachers, especially when compared to ELA teachers, while this trend is not readily apparent in teacher effectiveness as measured by EVAAS. This tends to suggest that, as currently written, student surveys do not appropriately reflect the instructional and pedagogical differences between math, science, and ELA classrooms. It is recommended that surveys become content specific, so as to accurately measure effective teaching in each individual content area.

CROSSWALK STUDENT SURVEY QUESTIONS WITH THE PRINCIPAL OBSERVATION RUBRIC TO CREATE A MORE COMPREHENSIVE EVALUATION PROCESS

Because student surveys have been shown to produce reliably consistent results and because these results can be tied to a teachers' EVAAS statuses, it is recommended that student survey questions be crosswalked with the principal observation rubric in order to create a more comprehensive evaluation and improvement process for teachers. By analyzing student survey results and aligning questions to the principal observation rubric, principals and school instructional leaders will be better able to offer more comprehensive feedback and support aimed at improving the instructional quality of teachers. An example of how this alignment could look is included in the appendix ***(see Table 5)***.

INCLUDE MULTIPLE DECISION CRITERIA IN TEACHER EVALUATIONS

The results of this study indicate that student surveys can explain about 20% of the variation seen in EVAAS scores. This suggests that, while comprehensive, EVAAS does not capture everything when it comes to measuring effective teaching. Additionally, while measuring student academic growth is an important metric in evaluating a teacher's effectiveness, it does not account for the non-cognitive and holistic benefits realized through a sound, basic education. By using multiple decision criteria to evaluate teachers (part of which, as suggested by this report, should be student surveys), the teacher evaluation process can move towards a more optimal position that not only improves outcomes for teachers, but for students as well.

APPENDIX

Table 1. Elementary Factor Loadings

The elementary factor loadings show the survey questions that loaded onto each of the three factors generated. These factors are used as a way to condense similar questions (i.e., questions that are scored similarly to each other) in order to reduce the number of variables used in the study.

Variable	Factor 1	Factor 2	Factor 3	Uniqueness
am_q110_5			0.5439	0.5404
am_q111_5	0.6451			0.3993
am_q112_5			0.8358	0.2685
am_q120_5	0.6579			0.5522
am_q126_5	0.7818			0.2719
am_q138_5		0.6530		0.4069
am_q143_5	0.5834			0.5758
am_q155_5		-0.8887		0.1885
am_q157_5	0.7011			0.4210
am_q165_5	0.7443			0.3583
am_q168_5				0.6166
am_q169_5	0.6894			0.5135
am_q176_5				0.6114
am_q185_5	0.6799			0.3845
am_q208_5	0.7004			0.3014
am_q209_5				0.6870
am_q210_5				0.8667
am_q212_5	-0.6663			0.3773
am_q213_5	0.7821			0.3221
am_q214_5	0.5833			0.5821
am_q219_5	0.8234			0.2305
am_q220_5		0.5522		0.4379
am_q29_5	0.5923	0.5105		0.3871
am_q38_5		0.8688		0.1936
am_q74_5	0.6868			0.5059
am_q83_5			0.8290	0.3097
am_q95_5	0.7408			0.2724

(blanks represent abs [loading] <.5)

Table 2. Secondary Factor Loadings

The secondary factor loadings show the survey questions that loaded onto each of the two factors generated. These factors are used as a way to condense similar questions (i.e., questions that are scored similarly to each other) in order to reduce the number of variables used in the study.

Variable	Factor 1	Factor 2	Uniqueness
m_b89	0.7893		0.3297
m_b141	-0.5220		0.6534
m_b44	0.7686		0.3479
m_b29	0.7749		0.3363
m_a10	0.6879		0.5157
m_b34	0.7673		0.3784
m_b146	0.6391		0.5811
m_b21			0.7610
m_b128	0.6556		0.5646
m_b36	0.6757		0.5041
m_b45	0.6531		0.5570
m_b136	-0.5434		0.6683
m_b1	0.7029		0.4997
m_b130	0.6760		0.5227
m_b17	0.7663		0.3903
m_b90	0.6668		0.5252
m_b80	0.8098		0.3219
m_b147	0.8189		0.3064
m_b135			0.7857
m_b129	0.5892		0.5953
m_b155	0.5056		0.6710
m_b145	0.6442		0.5664
m_b83	0.7106		0.4642
m_b58	0.6891		0.5052
m_b114		-0.5633	0.6666
m_b46		0.6610	0.4532
m_b6			0.6355
m_b49		0.6739	0.4129
m_b138		-0.7670	0.4013
m_b113		-0.7004	0.5094
m_b112		0.6016	0.5824

(blanks represent abs [loading] <.5)

Table 3. Descriptive Statistics of Variables for Elementary and Secondary Data

Variables	Elementary N = 1519	Secondary N = 1209
Teacher EVAAS Score	**.228 (-11.36-8.85)**	**.432 (-12.12-13.53)**
Teacher Effectiveness	**2.9 (1-5)**	**2.83 (1-5)**
Academic Support	-2.83E-10 (-5.95-2.56)	-4.42e-10 (-4.58-2.29)
Classroom Management	-6.71E-10 (-2.07-2.61)	-3.01e-10 (-2.7-2.94)
Academic Press	-3.03E-10 (-3.95-2.61)	N/A
% EDS Student	***56.14% (4%-98.8%)***	***52.73% (0-97.9%)***
% Minority Student	58.43% (5%-100%)	53.76% (1.2%-99.6%)
Teacher Absences	**9.89** (1-26.63)	**13.40** (.51-32.07)

Table 4. Multinomial Logistic Regression Marginal Effects

	Elementary			Secondary		
Variable	**Does Not Meet Expected Growth**	**Meets Expected Growth**	**Exceeds Expected Growth**	**Does Not Meet Expected Growth**	**Meets Expected Growth**	**Exceeds Expected Growth**
Academic Support	-.79%	.53%	.26%	**-3.7%*** **	.48%	**3.22%* **
Classroom Management	**-4.8%*** **	-.61%	**5.41%*** **	**-4.3%*** **	.31%	**4.03%** **
Academic Press	**-3.4%*** **	-1.99%	**5.40%*** **			
% EDS Student	**.19%*** **	-.10%	-.09%	**.28%*** **	**.36%*** **	**-.65%*** **
% Minority Student	-.06%	-.04%	.09%	**-.14%** **	**-.33%*** **	**.48%*** **
Grade	1.30%	**10.9%*** **	**9.64%*** **	-.10%	1.6%	-1.5%
Class Size	**.42%** **	-.34%	-.15%	.08%	-.41%	.32%
Teacher Absences	.50%	-.38%	.16%	.03%	-.40%	.37%
Math	.33%	-6.7%**	6.41%**	**13.9%*** **	**-30.5%*** **	**16.5%*** **
Science	2.40%	-5.2%*	2.81%	**20.36%*** **	**-40.2%*** **	**19.8%*** **
R^2	.15			.07		
F	5.48**			2.42*		

**p<0.1, **p<0.05, ***p<0.001*

Table 5. Elementary Factor Alignment

Factor	SS Question	Standard 5	Observation Rubric
1: Captivate	219: I like the ways that we learn things in this class. (captivate)	III: Teachers know the content they teach IV: Teachers facilitate learning for their students	IIIb: Teachers direct students' natural curiosity into content specific learning IIId: Teachers make content relevant to students IVc: Teachers use a variety of instructional methods
1: Care	126: My teacher in this class makes me feel that he/she really cares about me. (care)	I: Teachers demonstrate leadership II: Teachers establish a respectful environment for a diverse population of students	Ie: Teachers demonstrate high ethical standards through ensuring fair treatment for all students IIa: Teachers create a positive, nurturing environment; foster positive relationships with students IIc: Teachers treat students as individuals
1: Care	29: My teacher is nice to me when I ask questions.	II: Teachers establish a respectful environment for a diverse population of students	IIa: Teachers create a positive, nurturing environment; foster positive relationships with students IIc: Teachers treat students as individuals
1: Care	95: I like the way my teacher treats me when I need help.	I: Teachers demonstrate leadership II: Teachers establish a respectful environment for a diverse population of students	Ie: Teachers demonstrate high ethical standards through ensuring fair treatment for all students IIa: Teachers create a positive, nurturing environment; foster positive relationships with students IIc: Teachers treat students as individuals
1: Challenge	208: My teacher makes sure that I try to do my best.	I: Teachers demonstrate leadership II: Teachers establish a respectful environment for a diverse population of students	Ia: Teachers empower students to learn IIa: Teachers create a positive, nurturing environment
1: Clarify	165: My teacher checks to make sure we understand what he/she is teaching us.	IV: Teachers facilitate learning for their students	IVg: Teachers communicate effectively with students IVh: Teachers use a variety of methods to assess student learning

2: Captivate	120: We have interesting homework.	III: Teachers know the content they teach IV: Teachers facilitate learning for their students	IIId: Teachers make content relevant to students IVc: Teachers use a variety of instructional methods
2: Captivate	219: I like the ways that we learn things in this class.	I: Teachers demonstrate leadership III: Teachers know the content they teach IV: Teachers facilitate learning for their students	Ia: Teachers empower students to learn IIIb: Teachers know their content; direct students' natural curiosity into learning in the content area IVc: Teachers use a variety of instructional methods IVh: Teachers use a variety of methods to assess student learning
2: Captivate	74: School work is interesting.	III: Teachers know the content they teach IV: Teachers facilitate learning for their students	IIId: Teachers make content relevant to students IVc: Teachers use a variety of instructional methods
3: Challenge	112: My teacher pushes us to think hard about things we read.	I: Teachers demonstrate leadership II: Teachers establish a respectful environment for a diverse population of students IV: Teachers facilitate learning for their students	Ia: Teachers empower students to learn IIa: Teachers create a positive, nurturing environment IVe: Teachers promote critical thinking and problem solving skills
3: Challenge	83: My teacher pushes everybody to work hard.	I: Teachers demonstrate leadership II: Teachers establish a respectful environment for a diverse population of students IV: Teachers facilitate learning for their students	Ia: Teachers empower students to learn IIa: Teachers create a positive, nurturing environment IVe: Teachers promote critical thinking and problem solving skills
4: Control	155: Students behave so badly in this class that it slows down our learning.	I: Teachers demonstrate leadership IV: Teachers facilitate learning for their students	Ia: Teachers lead in their classrooms; they create safe, orderly environments IVg: Teachers communicate effectively
4: Control	38: My classmates behave the way my teacher wants them to.	I: Teachers demonstrate leadership IV: Teachers facilitate learning for their students	Ia: Teachers lead in their classrooms; they create safe, orderly environments IVg: Teachers communicate effectively

Works Cited

- *State of the Sates 2013 Connect the Dots: Using Evaluations of Teacher Effectiveness to Inform Policy and Practice.* Rep. National Council on Teacher Quality, Oct. 2013. Web.
- *Student Perception Surveys and Teacher Assessments.* Rep. Hanover Research, Feb. 2013. Web.

By Joe Maugeri, Julia Burrus Pierson, Vincent Reitano, and Qi Wang Xing

The Financial and Business Services Area is in its ninth year of the Research Intern Program. The Program is designed to help build a quality research program within NCDPI to supplement and supply data for discussions related to procedural, process, and policy changes. This year's program included students from Duke University's Master of Public Policy program, North Carolina State University's Master of Public Administration program, and The University of North Carolina at Chapel Hill's Master of Public Administration and Doctorate in Education programs. The intern program is managed by Eric Moore (919-807-3731) and Kayla Siler (919-807-3824) | intern_research@dpi.nc.gov

NC DEPARTMENT OF PUBLIC INSTRUCTION **: : June St. Clair Atkinson, Ed.D., State Superintendent : : 301 N. Wilmington Street : : Raleigh, NC 27601-2825**
In compliance with federal law, the NC Department of Public Instruction administers all state-operated educational programs, employment activities and admissions without discrimination because of race, religion, national or ethnic origin, color, age, military service, disability, or gender, except where exemption is appropriate and allowed by law.

Inquiries or complaints regarding discrimination issues should be directed to: Dr. Rebecca Garland, Deputy State Superintendent
6368 Mail Service Center, Raleigh, NC 27699-6368 :: Telephone: (919) 807-3200 :: Fax: (919) 807-3388