BEST IN CLASS: AN INVESTIGATION OF THE RELATIONSHIP BETWEEN TEACHER QUALITY INDICATORS AND STUDENT PERFORMANCE IN NORTH CAROLINA

Jason M. Sinquefield

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Education in the Educational Leadership Program in the School of Education.

Chapel Hill
2017

Approved by:

Kathleen Brown

Brian Gibbs

Eric Houck

Christopher Scott

Dana Thompson-Dorsey

ABSTRACT

Jason M. Sinquefield: Best in Class: An Investigation of the Relationship Between Teacher
Quality Indicators and Student Performance in North Carolina
(Under the direction of Eric Houck)

The purpose of this study was to investigate the relationship between North Carolina's common
teacher quality indicators and student performance. To that end, this study sought to determine in
what way North Carolina's teacher quality indicators significantly predicted student growth as
measured by EVAAS scores, as well as the profile of the North Carolina teacher whose students
exceeded expected growth. Spence's (1974) Signal Theory was used to explain the dissonance
between what is used by educators (i.e. years of experience, master's degree, and North Carolina
Teacher Evaluation Rubric rating) and the state (EVAAS status) to demonstrate high quality
teaching.

The study used secondary data sets containing North Carolina teacher quality and school
and teacher demographic data furnished by local, state, and national agencies for the 2010-2011
through 2013-2014 school years, totaling more than 101,000 teacher observations. Multiple
regression analyses were run to investigate relationships using teacher and school variables and
informed answers pertaining to the research questions of the study. A number of controls for
teacher, school, and district effects were included in regression models.

Results showed a significant positive relationship for teacher years of experience and
North Carolina Teacher Evaluation Rubric ratings with EVAAS scores. Higher school non-
White populations consistently yielded a significant negative relationship with EVAAS scores.
Lastly, teachers who were White, male, had higher North Carolina Teacher Evaluation Rubric

ratings, and taught at non-Title 1 schools with smaller minority student populations were found to have the highest probability of exceeding expected growth.

In addition to connecting North Carolina to existing research on the efficacy of teacher quality indicators in predicting student performance, this study also offers information to educational leaders and policy makers interested in improving teacher recruitment, compensation, and retention strategies. However, this study also raises concerns on the existing system of teacher evaluation in North Carolina, as the comprehensive models used explained no more than five percent of the variation in EVAAS scores. Similarly, demographic characteristics of teachers and schools had a significant impact on student performance that will require further research and investigation in order to improve the explanatory power of the findings presented.

ACKNOWLEDGEMENTS

This achievement is also your achievement.

To my wonderful wife, Adrienne: thank you for your unconditional support, your grace in accepting late nights and early mornings when schoolwork called, your willingness to hold me accountable for deadlines and action plans I set, and for being my rock during the times when I doubted myself. Most importantly, I thank you for encouraging me to dream and dream big; this is only the beginning.

To my daughter, Vera: thank you for making those late nights and early mornings—since you were usually awake then too—that much more beautiful. I cannot count the number of times your smile, laugh, new word, or singing and dancing gave me the boost I needed to read that extra article or write that extra page. May the curiosity, wonder, and passion for new knowledge that drove me to undertake this journey stay with you forever.

To my daughter, Laurel: thank you for further focusing me on what is most important. I pray that you successfully fight for each day, and everything you believe in, the way you have these first few months of your life.

To the God I love and my Lord and Savior, Jesus Christ: thank you for the opportunity to use this moment to glorify You and the everlasting love You provide. Your wisdom, guidance, and grace gave me the courage and skills I needed to cross the finish line. This is for You.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

In accordance with an ever-changing political and economic landscape, considerable attention and discussion continues to center around the fundamental principles guiding teacher compensation. For decades, teachers have been compensated according to a single salary schedule, whereby teachers are rewarded by years of experience and educational attainment (Omps, 2011-2012). There are some that feel this is the fairest way to compensate teachers, while others believe that merit-based approaches, like those found in the private sector, are more appropriate (Edmonds, 2012; Malin, 2009). The resulting arena, however, still relies on a system of evaluation and compensation that aims to promote equality and not equity; that is, finding a variable or set variables—like teacher experience or student test scores—that deliver a consistent teacher quality message across all contexts.

The ongoing focus on teacher compensation emanates from evidence regarding the underperformance of American students on national and international tests (National Commission on Educational Excellence [NCEE], 1983; Peterson, Lastra-Anadon, Hanushek, & Woessmann, 2011). Because student performance is naturally correlated to the quality of instruction received, policymakers are looking for ways to systemically reward outstanding teachers and attract high-potential individuals into the classroom (Feldman, 2000). As a result, increasing the number of high quality teachers in United States classrooms, particularly in areas that serve traditionally low-performing students, has become a key component in most education reform movements (Edmonds, 2012). Thus, the reevaluation and reconstruction of traditional teacher compensation and evaluation models continues to have a place in education policy

discourse, as teaching has garnered a reputation as a profession of relatively stagnant pay and limited career upward mobility despite performance (Omps, 2011-2012). These conditions have led to competent teachers leaving the classroom, reformers have argued, with few quality replacements willing to enter the profession (Bagdon, 1985).

Statement of the Problem

The guiding principle behind performance pay plans is the notion that employment compensation should be tied to the level of excellence achieved by the employee, with higher performance leading to a higher salary (Citron, 1985). In aligning teacher pay with this ideology, districts would be employing a compensation principle utilized in many other professions, whereby hard work and improved skills that led to measurable successes entitled the employee to higher earnings (Calhoun & Protheroe, 1983). While it is believed that both experience and educational level correlate with improved skills and performance, research has not been able to fully support this claim (Omps, 2011-2012). As a result, compensation systems that rely solely on these factors have shown to be both inefficient and inadequate (Goldhaber, Dearmond, & Deburgomaster, 2011; Malin, 2009; Omps, 2011-2012). In search of a scalable model that captures all components of quality, reformers believe that allowing at least a portion of teacher compensation to be derived from performance indicators incentivizes professional growth and exploration, and therefore rewards teacher excellence (Feldman, 2000; Harrison & Cohen-Vogel, 2012; Springer, 2009). This position is held by reformers despite a dearth of evidence that such a "kitchen sink" approach improves student outcomes and without addressing underlying concerns related to inequity in education and teacher evaluation systems.

A local example of extending this ideology to teacher compensation policy, North Carolina Governor Pat McCrory announced the creation of the Career Pathways Fund in 2014,

allowing for and supporting North Carolina school districts interested in piloting locally created

teacher performance pay plans (State of North Carolina, Office of the Governor, 2014). The

Fund supplied initial and ongoing financial support to these plans, provided districts received

approval from the North Carolina State Board of Education and created plans that were pursuant

to the Fund's stated objectives. School districts were required to match all Fund disbursements

received. Later a part of the 2014-2015 budget passed by the North Carolina General Assembly

(Appropriations Act of 2014, 2014), a number of districts submitted teacher compensation

models that reflected both local strategic plans and state objectives. One key Career Pathways

Fund objective, however, garnered significant attention and scrutiny: rewarding excellent

teaching. When coupled with the state's embracement of value-added data (Dewitt, 2014), this

objective challenged districts to create plans that accurately identified and rewarded teachers

whose students performed well on state tests.

     The decentralized nature of the Career Pathways Fund program created an opportunity

for school districts to unknowingly use teacher quality indicators that may not correlate with

excellent teaching, manifested in high student performance in value-added data. In fact, state

lawmakers expressed concerns with the quality and creativity of the 75 plans submitted for

approval in the first year of the program (Ball, 2015). The disconnect can in part be explained by

a lack of published research demonstrating the relationship between the state's most popular

teacher quality indicators—master's degrees, the North Carolina Teacher Evaluation Rubric

ratings, and years of experience—and its preferred value-added metric, EVAAS. EVAAS, North

Carolina's metric for student performance, describes students and teachers as having not met,

met, or exceeded expected growth. Growth is determined in this model by using the student's

own performance on previous tests and other demographic information to estimate the amount of expected growth for that student during one school year.

This study informs educational leaders interested in creating differentiated teacher compensation plans of the indicators that most accurately predict high student growth, while shedding light on ways that existing inequities in the system related to race, class, and gender influence the predictive power of those indicators. To that end and from the perspective of Spence's (1974) Signal Theory, the researcher examined the efficacy of North Carolina's stated teacher quality indicators in predicting its desired EVAAS student performance outcome: Exceeds Expected Growth.

Purpose of the Study and Research Questions

The purpose of this study was to investigate the relationship between North Carolina's teacher quality indicators—years of experience, master's degree attainment, North Carolina Teacher Evaluation Rubric ratings—and student performance. The following research questions guided the study:

1. In what way, if any, do North Carolina's teacher quality indicators significantly predict student growth, as measured by EVAAS?

2. What is the profile of the North Carolina teacher whose students exceed expected growth?

Significance of the Research

The rise in popularity of performance-based pay systems originated with the release of the *A Nation at Risk* report during the Reagan Administration, which called for increased attention to and an evaluation of the nation's education system (NCEE, 1983; Omps, 2011-2012). Among the key themes of the report was a recommendation that teachers be paid

according to their performance in the classroom (NCEE, 1983). As a result, performance pay for teachers was one of the major initiatives promoted by the Reagan Administration's New Federalism plan (Goodman & Melia, 1988-1989; Poltrock, 1984). In voicing his fervent support of teacher merit pay, President Reagan (1985) stated:

> Today, America boasts thousands of fine teachers, but in too many cases teaching has become a resting place for the unmotivated and the unqualified. And this we can no longer allow. We must give our teachers greater honor and respect…and we must pay and promote our teachers according to merit (p. 3-4).

With the teacher excellence movement prominently positioned in national discourse in the early 1980s, performance pay established itself as a part of education reform movements for years to come (Veir & Dagley, 2002). Although President Reagan's promotion of merit pay made the compensation system a popular component of a number of education reform proposals across the country, performance-based pay did not fully catch on during his time in office (Goodman & Melia, 1988-1989). However, merit pay systems would continue to be considered and proposed as a practical tool for both recruiting and retaining quality teachers (Springer, 2009). The concept has also been consistently considered as one of the many remedies needed to fix the United States education system (Omps, 2011-2012).

After nearly two decades of discussion and stagnation in student performance, policymakers again began to call for merit pay principles in their teacher compensation plans (*The Friday Report*, 2005). Citing the inflexibility of the single salary schedule in incentivizing innovation and excellence, state leaders began to explore creative ways to reward teachers who performed well, particularly those in less than ideal circumstances. Most of the recent merit pay proposals, beginning in the mid-2000s, focused on providing bonuses for teachers in hard to staff schools and subjects in addition to all those who used professional growth opportunities to significantly raise student performance (Superfine, 2011).

Representative of the increase in popularity of merit pay plans for teachers in 2006, the Department of Education created the Teacher Incentive Fund, which provided support to districts across the country that were interested in instituting merit pay compensation plans (Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 2006). Later, President Obama would support the concept of performance pay in his own education reform initiatives in the American Recovery and Reinvestment Act (American Recovery and Reinvestment Act of 2009, 2009; Daniel & Dyson, 2009-2010). Coinciding with this rise in popularity and support, several states passed legislation to aid in the creation of performance-based compensation measures for teachers (Ariz. Rev. Stat. Ann. § 15-977, 2007; Colo. Rev. Stat. Ann. § 22-63-401, 2014; Minn. Stat. § 122A.60, 2013; Miss. Code Ann. § 37-19-7, 2013; S.C. Code Ann. § 59-18-1500, 2009; Tex. Educ. Code Ann § 21.701-705, 2009).

By evaluating the efficacy of teacher quality indicators on predicting student performance, this study informs educational leaders interested in identifying excellent teaching in North Carolina. Such a study is significant, as little published research has been done examining the relationship between North Carolina's teacher quality indicators and its EVAAS value-added model. Likewise, noting the impact of current systemic inequities should contribute to more just and honest policy analysis and construction in the future. The conceptual framework used, described next, has also rarely been used in the field of educational leadership and is therefore a new approach to teacher quality research.

Overview of Methodology & Conceptual Framework

This study used Spence's (1974) Signal Theory to explain the dissonance between what is used by educators (years of experience, master's degree, and North Carolina Teacher Evaluation Rubric rating) and the state (EVAAS status) to demonstrate high quality teaching. To

6

that end, Signal Theory establishes a framework for describing how employers and employees communicate qualitative characteristics (i.e. excellent teaching) through a diversified set of signals. A lens for evaluating and discussing the implications of the quality of those signals is also made available through this approach.

This correlational study—the most logical approach given the researcher's intent on investigating relationships between variables—requires the use of teacher quality data aggregated by North Carolina pursuant to compliance with the Race to the Top (RttT) initiative. To that end, teacher quality indicators years of experience, educational attainment, and North Carolina Teacher Evaluation Rubric ratings will be used to determine relative predictability of EVAAS status using regression analysis. Regression analysis will also be used to predict variability in Exceeds Expected Growth teachers in North Carolina.

The sample used in this study was comprised of full-time North Carolina public school teachers that: taught during the 2010-2011 through 2013-2014 school years, received a North Carolina Teacher Evaluation Rubric rating, and received an EVAAS score. These criteria were chosen in order to ensure representative data existed for the teachers used in the statistical procedures and analyses of the study.

Assumptions

The primary assumption of this study is that quality teaching is a phenomenon that can be finitely defined and explained. For decades, educational leaders and policymakers have tried to both define and incentivize excellent teaching, with varied results. To that end, they have created a system that leverages these assumptions in the interest of evaluating teachers equally on indicators that, collectively, indicate "merit" without significant effort to control for local, regional, and statewide context. This study, contributing to the literature on the efficacy of

teacher quality indicators on identifying excellent teaching, does so while assuming that excellent teaching is quantifiable and defined using North Carolina's evaluation model. Similarly, this study assumes that teacher performance is partly defined by EVAAS scores. This was done due to the State of North Carolina's decision to use the metric as such; EVAAS growth status is a part of teacher ratings in North Carolina, commonly referred to as Standard 6 on the North Carolina Teacher Evaluation Rubric.

The teacher quality indicators used—years of experience, master's degrees, and North Carolina Teacher Evaluation Rubric ratings—were used to ascertain alignment with the state's longtime use of them as proxies for teacher quality. Similarly, the researcher only used assessments likewise used by the North Carolina, like End of Course (EOC) and End of Grade tests to determine EVAAS status; other performance measures, like ACT or SAT scores, were therefore not used.

Limitations

A primary limitation of this study relates to the variability within teacher quality indicators. Not all master's degrees received by teachers are created equally or even have consistent application to classroom practices (e.g. master's degree in Mathematics versus Mathematics Education), and the programmatic rigor and professional growth attained by graduates is difficult to control or correct for in this research design. Thus, conclusions drawn about the efficacy of master's degrees present generalizability concerns; however, the practice of treating all master's degrees, whether they be content area or pedagogically focused, is consistent with the state of North Carolina's treatment of them. Similarly, principals and assistant principals deliver teacher ratings on the North Carolina Teacher Evaluation Rubric, which creates concerns for consistency in application that the researcher cannot control for in this study. However, these

ratings represent some of the best information available to the researcher in the area of excellent teaching and are in fact used by the state already in determining teacher quality. Further, the data set used by the researcher did not include National Board-Certified Teachers (NBCT). However, this certification is a national measure of teacher quality—as opposed to the state-developed North Carolina Teacher Evaluation Rubric— and, like the Rubric, evaluates teachers based on observable in-class behaviors. Lastly, it is important to note that assuming North Carolina's model as the foundation for this study also forced the researcher to take on the variability caused by and the influence of inequities in the system related to race, class, and gender on the findings presented.

Another limitation is the use of EVAAS and Exceeds Expected Growth as metrics for teacher excellence. An extension of one of the assumptions previously presented, the use of this metric also creates generalizability concerns for this research as not all states use EVAAS to determine effective teaching. This limitation is mitigated by the intended scope of the research; this study intends to inform North Carolina's educational leaders of the efficacy of current teacher quality indicators in predicting excellent teaching. Lastly, there are teachers in North Carolina with missing or incomplete data on the indicators studied; this reality is a limitation, but is mitigated due to the number of observations and methodological techniques used in the study.

Definitions of Key Terms

Educational Attainment – describes the degree level received by a teacher: bachelor's degree, master's degree, or advanced degree.

Educational Value-added Assessment System (EVAAS) - the value-added model used by the state of North Carolina to determine if a student did not meet, met, or exceeded expected student growth. The consolidation of all student results determines teachers' ratings on this scale.

Master's Degrees – any graduate school degree, whether it is in a content area (e.g. Mathematics) or in Education (e.g. Mathematics Education), attained by a teacher who teaches in North Carolina.

Performance Pay - Performance-related pay (e.g. performance pay, merit pay) is a compensation approach that rewards employees in part for the educational outputs they generate rather than solely for the skills and knowledge they input. In the education profession, this means financially rewarding teachers for indicators other than or in addition to years of experience or advanced degree attainment (Liang & Akiba, 2011).

Signal Theory – used to describe "behavior when two parties (individuals or organizations) have access to different information" (Connelly, Certo, Ireland, & Reutzel, 2011, p. 39).

Student Growth – determined based on comparing the student's subsequent performance to his/her previous performance on a standardized test.

Student Proficiency – determined by his/her performance relative to a threshold score on a standardized test that indicates mastery of content.

Teacher – a person instructing students in North Carolina public schools in the 2013-2014 school year

Teacher Compensation Model – a system for paying teachers based on a variety of characteristics and/or outcomes

Teacher Quality Indicators – characteristics used to describe teacher competency

Years of experience – the number of years a teacher has taught in the public-school system in North Carolina

CHAPTER II: LITERATURE REVIEW

The focus of this study was to investigate the efficacy of North Carolina's most used teacher quality indicators in predicting student performance. To that end, this literature review begins by presenting and arguing for Signal Theory as the conceptual framework of this study. A review of the literature on the major teacher quality indicators—years of experience, master's degrees, the North Carolina Teacher Evaluation Rubric, and value-added models—will follow. These indicators were chosen due to their historical significance in the teacher quality literature and their role in current North Carolina teacher quality and compensation policy discourse. The chapter concludes with a summary of major points presented.

Conceptual Framework: Spence's Signal Theory

This investigation of the efficacy of North Carolina's most used teacher quality indicators in predicting high student performance leveraged Spence's Signal Theory. Conceptualized in his later-published doctoral dissertation, Spence's (1973) Signal Theory is used to explain the behavior of different groups who do not share the same information. In this model, one group—the sender—attempts to signal information to another group—the receiver—who then must determine the meaning and/or value of the aforementioned signal. The information signaled by the sender is typically a declaration of overall quality that the receiver would therefore find valuable in evaluating the sender.

Figure 1 provides a visual representation of the signaling cycle that Spence presents. The signaler has an underlying quality that needs to be demonstrated to the receiver in the larger signaling environment. To that end, the signaler communicates to the receiver via a signal (e.g.

higher education attainment) that serves as a proxy for underlying quality. The receiver must then determine if the signal does in-fact project the underlying quality the signaler wishes to portray, an evaluation by the receiver that generates feedback of signal efficacy to both the signaler and the larger signaling environment.

Figure 1

Signaling Timeline
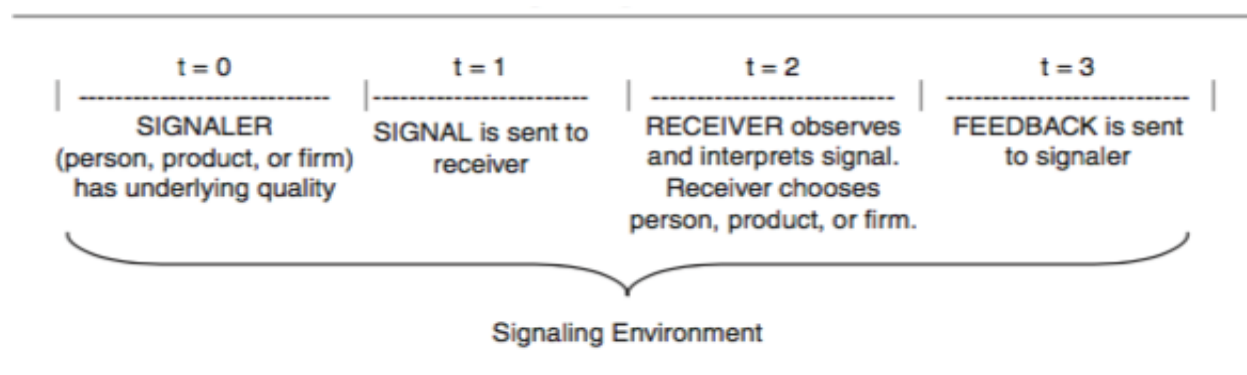
Source: Connelly, Certo, Ireland, and Reutzel, 2011, p. 44

Note: t = time

The genealogical application of this theory, as it was positioned in Spence's first publication, was used in describing the interplay of actors in the labor market: potential employees and prospective employers (Connelly, Certo, Ireland, and Reutzel, 2011). To that end, the model described the ways potential employees tried to differentiate themselves from others in the labor pool to prospective employers, namely through higher education (Spence, 1973). However, in the labor market, where numerous senders have access to many of the same signals (e.g. higher education), Spence posited that the signaling equilibrium, where certain signals are believed by the receiver to equate to quality of the sender and are therefore increasingly signaled by senders, can become disrupted after the finite number signals available no longer differentiate sender quality due to their pervasiveness in the market. This is especially true for signals like

higher education and years of experience, which are pervasive in education labor markets today and vary considerably in quality despite equivalences in the names of degrees earned and time served, respectively. What results from signal equilibrium disruption is what Stiglitz (2002) described as information asymmetry, whereby signaler(s) and receiver(s) have access to different information. This information includes public information, such as degrees attained and years of experience, as well as private information, such as the true quality of the aforementioned degrees and experiences, that is only available to some.

Reducing information asymmetry, then, is both the principal focus of Signal Theory and its primary application in the decades since its inception (Spence, 2002; Stiglitz, 2002). In addition to the labor market applications as part of Spence's original work, scholars from a number of disciplines have used Signal Theory to explain the behavior between groups with diversified sets of information. Table 1 provides a visual summary of this relevant Signal Theory literature investigating the relationships between signals in various fields.

Table 1

Summary of Literature Review: Signal Theory Across Disciplines

Field	Author(s)	Year	Signal	Findings
Management	Lampel & Shamsie	2000	Marketing Strategy	Strong relationship between pre-conceived value of a film and subsequent marketing investment
Management	Miller & Triana	2009	Board Diversity	Positive relationship between board diversity and firm's reputation
Management	Zhang & Wiersma	2009	Certified Financial Statements	Reliability of statements strongly correlated to CEO attributes
Entrepreneurship	Filatotchev & Bishop	2002	Board Composition	Strong relationship between board composition (e.g. experience) and initial valuation
Entrepreneurship	Busenitz, Fiet, & Moesel	2005	Degree of Founder Investment	No relationship between degree of New Venture Team (NFT) firm investment and success in acquiring VC funding

Table 1 (*cont.*)

Field	Author(s)	Year	Signal	Findings
Entrepreneurship	Lester, Certo, Dalton, Dalton, & Cannella	2006	Board Composition	Strong relationship between Top Management Team (TMT) and initial valuation
Organizational Behavior Management & Human Resources	Srivastava	2001	Offers & Counter-offers	Offers and counter-offers reflected bargaining position information not available to receiver
Organizational Behavior Management & Human Resources	Hochwater, Ferris, Zinko, Arnell, & James	2007	Political Influence	Strong relationship between employee influence and performance reputation
Organizational Behavior Management & Human Resources	Casper, Wayne, & Manegold	2013	Company Culture Data	Recruitment efforts more effective when aligned to values (e.g. inclusivity) rather than demographic information

In the management field, for example, Zhang and Wiersema (2009) used Signal Theory to explain the strong positive correlation between both the number of shares held by a CEO and his/her activity on other company boards and the reliability of financial statements that had been certified by the CEO. In this study, 419 CEOs attempted to signal their firm's underlying quality to investors through certified financial statements; however, the quality and reliability of those signals were found to be strongly associated with other CEO attributes. Miller and Triana (2009) also used Signal Theory in their management study to explain the tendency of top performing firms to seek out corporate board diversity. In their study of 326 *Fortune 500* companies, they found a positive relationship between the racial diversity of a board and the firm's larger reputation. Board diversity, then, served as a signal to investors about the firm's underlying quality and leadership capacity for growth in complex markets. Similarly, Lampel & Shamsie

(2000) used Signal Theory to explain the pre-movie release strategies of studios marketing a new film. In their study of 409 feature films, it was found that a strong relationship existed between studios' pre-conceived value of the feature film—that is, whether it exhibited a strong or weak signal to the market—and the subsequent marketing strategy (e.g. total marketing budget, distribution strategy) executed.

The field of Entrepreneurship has similarly leveraged Signal Theory to explain the behaviors of its actors—particularly during Initial Public Offerings (IPOs)—as most firms navigating the IPO process are relatively new to the market and therefore struggle to demonstrate long-term investment value (Certo, 2003). To that end, Filatotchev & Bishop (2002), in a study of 251 IPOs in the London Stock Exchange (LSE), found that a strong relationship existed between board composition and the likelihood that investors would undervalue initial stock purchase prices. Board composition, wherein members had varying degrees of experience and industry status, therefore served as a signal for an individual company's potential as an investment. Lester, Certo, Dalton, Dalton, and Cannella (2006) had similar findings, as their study of 209 IPO prospectuses found that a strong relationship existed between the perceived strength of the "Top Management Team (TMT)" (p. 2) and the valuation eventually given to firms. In this study, board composition again served as a signal to potential investors of investment worthiness. Applications of Signal Theory have extended into the venture capital strand of the Entrepreneurship literature as well, as researchers have aimed to explain the investment decisions of venture capital firms. For example, Busenitz, Fiet, and Moesel (2005), in their study of 183 venture capital investments monitored over a 10-year period, found no significant relationship between the degree to which a New Venture Team (NVT) was invested in their own venture and the eventual performance of the venture. This study aimed to clarify the

informational asymmetries caused by the "value and commitment" signals (p. 1) sent by NVTs—demonstrated by the degree to which the members were invested in the venture themselves— to venture capital firms in need of determining long-term investment potential.

Organizational Behavior Management and Human Resource Studies (OBM/HR) have likewise applied Signal Theory to explain the behaviors of current and potential employees. For example, Hochwater, Ferris, Zinko, Arnell, & James (2007) investigated the impact of an employee's ability to influence coworkers on the aforementioned employee's reputation. In this two-part study of 581 employees from a wide range of occupations, a strong relationship was found between the ability of an employee to influence coworkers and performance reputation. Thus, employees' ability to influence others served as an effective signal to fellow coworkers of underlying employee quality. Srivastava (2001) applied Signal Theory in order to investigate the effect of signals on the bargaining process. In this two-part study of 20 MBA students, it was found that actors in the bargaining process consistently presented counteroffers that reflected the underlying information received from a previously received offer. In this instance, initial and counter offers sent signals to the recipient regarding information about bargaining position not publicly available to the receiver. Recipients of offers and counteroffers, then, were tasked with both assessing the credibility of signals received through these offers and providing feedback (i.e. a counter signal) to the sender via a new counter offer. Casper, Wayne, and Manegold (2013) leveraged Signal Theory to explain the impact of signals sent by an employer about company culture on prospective employees' interest in employment. In this study of 300 employed adults, it was found that prospective employee beliefs and value systems, or "deep level differences" (p. 322), were more effective indicators of employment interest than indicators more demographic in nature. Recruitment efforts by employers, then, were shown to be more

effective when signals were aligned with the underlying values associated with company culture rather than demographic characteristics to that end (e.g. inclusivity vs. race/ethnicity).

The wide-ranging applicability of Spence's theory forms the basis of its critique, in that each new application only serves to further obscure the central purpose and foundation of the theory itself. This reality is most clearly demonstrated by the volume of research wherein Signal Theory is the conceptual lens despite a dearth of cross-discipline consensus around its tenets (Connelly, Certo, Ireland, & Reutzel, 2011). Thus, despite its rise in popularity and application to disciplines as far-reaching as anthropology (Bird & Smith, 2005), the broad appeal of Signal Theory has led critics to believe that it is insufficient as a standalone theory in their disciplines (Ehrhart & Ziegert, 2005; Highhouse, Thornbury, & Little, 2007). Whether lacking the ability to explain the relative power between signals in the signaling environment (Ehrhart & Ziegert, 2005) or the depth to account for the contextual nuance of signals sent (Highhouse, Thornubry, & Little, 2007), critics argue that Signal Theory requires adaptation and/or supplementation with other theories in order to more comprehensively describe phenomena in their disciplines (Highhouse & Hoffman, 2001; Lievens & Highhouse, 2003; Slaughter, Zickar, Highhouse, & Mohr, 2004).

The use of Signal Theory in a study in educational leadership is both a natural application considering previous uses in other disciplines and representative of the critique described above. However, the researcher believes that this framework is most suited for this study since the purpose of the work is to investigate the ability of teacher quality indicators to meaningfully predict eventual teacher excellence. This study is focused on informational symmetries and asymmetries between signalers (i.e. teachers) and receivers (i.e. educational leaders making hiring decisions), and is representative of the signaling equilibrium in Spence's

seminal work. Thus, because proper identification is the central intent, competing theories that focus on incentives (Jensen & Mecklin, 1976; Ross, 1973) are not applicable to this work. This intent also rules out production functions, a popular conceptual lens in education research, as this study does not assign differentiated costs to teacher quality indicators as a means for improving teacher productivity (Fortune & O'Neil, 1994; Hanushek, 1986; Hedges, Laine, & Greenwald, 1994; Knoeppel & Rinehart, 2008; Monk, 1989); rather, this study simply aims to identify which indicators are most likely to predict the desired outcome of teacher excellence.

This section introduced the conceptual lens, Signal Theory, which the researcher used in the investigation of the efficacy of North Carolina's most used teacher quality indicators in predicting student performance. The researcher used this conceptual lens, as it was in Spence's seminal work and in many disciplines since, as a means to reduce the informational asymmetries between available teacher quality indicators used to signal underlying quality. This section highlighted a number of studies available in the literature leveraging Signal Theory and also presented alternative conceptual lenses available to the researcher. The next section(s) will transition to a review of the literature on the major teacher quality indicators and metrics (i.e. signals) used by current and prospective teachers and education agencies to signal underlying quality in North Carolina: years of experience, master's degrees, the North Carolina Teacher Evaluation Rubric, and EVAAS.

Indicators in the Making: A Brief History

The genesis of teacher quality indicators is similar to that of other professions in the early 20th century: they existed for identifying and compensating the best performers (Calhoun & Protheroe, 1983; Springer & Gardner, 2010). However, early systems of identifying and compensating teacher quality were rife with inconsistency and discrimination based on race and

gender and, as a result, a movement began in the 1920's to end such practices and create salary equality for all similarly qualified teachers (Kelley, 1997; Omps, 2011-2012, Springer & Gardner, 2010). By 1950, single salary schedules had become the standard teacher quality and compensation system for most states (Calhoun & Protheroe, 1983; Podgursky & Springer, 2007). The single salary schedule was intended to protect against the discrimination of teachers who had similar skills and attributes, and teachers were compensated based on their length of service as well as the educational credentials they received rather than immutable characteristics like race and gender (Kelley, 1997; Omps, 2011-2012; Springer & Gardner, 2010). Grounded in the notion that more years of experience and education would result in improved teaching, this compensation system awarded improved teacher quality via incremental salary increases for length of teaching service and the amount and kind of educational attainment received, increasingly master's degrees (Feldman, 2000; Kelley, 1997; Omps, 2011-2012; Springer & Gardner, 2010). This rigid, consistent approach to teacher quality thus attempted to ensure that teachers were identified and paid equally if they were equally qualified. While the single salary schedule protected teachers from discrimination, it also began to limit the ability of districts to identify and reward teachers for excellence not identified in the schedule—rewarding all teachers the same based on experience and educational attainment, even those deemed ineffective, increasingly became a measure viewed as both costly and unfair (Brodsky, DeCesare, & Kramer-Wine, 2010; Omps, 2011-2012).

In the decades since the invention of the single salary schedule, a number of new approaches—namely, teacher observations and value-added models (VAMs)— have been introduced by state and local education agencies to identify and reward teacher quality. The use of teacher observations as a quality indicator rose in popularity in the late 1980s and has been

used in some capacity to identify teacher quality since (Citron, 1985; Edmonds, 2012; Figlio & Kenny, 2007; Rockoff & Speroni, 2011; Roye, 2010). Further, the use of standardized teacher observation instruments to identify and reward teacher excellence has grown in popularity of late, as evidenced by its inclusion in the North Carolina teacher evaluation process (NC State Board of Education, 2012) and, therefore, this study.

However, because of the subjectivity concerns inherent in teacher evaluation via in-school observation, the field of education has actively searched for more objective ways to provide school and teacher accountability in recent years by attempting to place more emphasis on student outcomes (Kelley, 2007; Podgursky & Springer, 2007). As a result, the use of value-added models (VAMs) has increased, as the prospect of measuring students based on growth, rather than just proficiency, has gained popularity (Amrein-Beardsley, 2013; National Council on Teacher Quality, 2011; Podgursky & Springer, 2007; Sanders & Rivers, 1996; Sanders, Saxton, & Horn, 1997). North Carolina has been a leader in the use of VAMs to identify and reward teacher quality through its state-wide use of VAM product SAS EVAAS (U.S. Department of Education, 2006), and is poised to continue its efforts in this area (Collins & Amrein-Beardsley, 2014; Dewitt, 2014).

This section provided a brief history of the most commonly used teacher quality indicators. Each indicator—years of experience, master's degrees, the North Carolina Teacher Evaluation Rubric ratings (i.e. teacher observations), and EVAAS—continues to be used by teachers and educational leaders to signal or determine underlying quality. For these reasons, each are included in this study as either an employee or employer signal. The next section(s) will provide a more in-depth review of each of the aforementioned signals as they relate to this study,

as well as present relevant literature on the ability of each indicator to influence or predict student performance.

Employee Signal 1: Teacher Years of Experience

Years of experience has long been considered an indicator of teacher quality, as evidenced by its continual inclusion in the single salary schedule (Kelley, 1997; Omps, 2011-2012, Springer & Gardner, 2010). The relevance of the indicator to this study is strong since length of service is still a metric that informs quality, compensation, and employment (e.g. tenure) designations in North Carolina (NC Department of Public Instruction, 2015; NC State Board of Education, 2012).

As an indicator, the experience indicator presents numerous advantages, namely that increasing years of service is achievable, unbiased, understandable, and popular with teachers (Goldhaber, Darmond, & DeBurgomaster, 2011). However, opponents of the years of experience teacher quality indicator argue that it does not align with improving student achievement (Brodsky, DeCesare, & Kramer-Wine, 2010; Omps, 2011-20112). This criticism is centered on the notion that all teachers with equal years of service are not equally excellent, meaning that years of experience as a teacher quality indicator does not differentiate between teachers' ability (Brodsky, DeCesare, & Kramer-Wine, 2010). Similarly, as a part of a single salary schedule that only identifies excellence and compensation through years of experience and educational attainment (discussed later), little incentive is given to teachers to improve instruction or accept difficult teaching assignments (Edmonds, 2012; Goldhaber, Dearmond, & Deburgomaster, 2011). Thus, opponents argue that the years of experience teacher quality indicator does not directly promote teacher excellence and, when combined with the rigidity and limited mobility of the single salary schedule, makes the profession seem less attractive to high potential prospective

teachers (Omps, 2011-2012; Podgursky & Springer, 2007). These criticisms and stated potential

externalities aside, the most salient criticism of the years of experience teacher quality indicator

is that it has not conclusively shown to impact student achievement (Goldhaber, 2002;

Hanushek, Kain, O'Brien, & Rivkin, 2005; Harris & Sass, 2011; Malin, 2009; Omps, 2011-2012;

Podgursky & Springer, 2007). Table 2 provides a visual summary of the relevant literature

investigating the strength of the relationship between teacher years of experience and student

achievement.

Table 2

Summary of Literature Review: Teacher Years of Experience

Author(s)	Year(s)	Indicator	Findings
Hanushek	1986, 1989, 1991, 1997	Weak	Consistently found to be a weak indicator of student achievement, at best
Greenwald, Hedges, & Laine	1996	Strong	Teacher experience and student achievement were related after re-analysis of Hanushek's work using advanced meta-analysis techniques
Dewey, Husted, & Kenny	2000	Strong	Teacher experience and student achievement were related after re-analysis of Hanushek's work using advanced meta-analysis techniques
Nye, Konstantopoulos, & Hedges	2004	Strong	Significant positive effects for teacher experience on 2nd grade reading and 3rd grade mathematics achievement
Rivkin, Hanushek, & Kain	2005	Weak	No relationship between teacher experience and student performance after the third year
Clotfelter, Ladd, & Vigdor	2006	Weak	Half of significant standard deviation increase shown by students with experienced teachers explained by effects of initial and second year teachers
Aaronson, Barrow, & Sanders	2007	Weak	Years of experience at best explained 8% of variation in student achievement
Clotfelter, Ladd, & Vigdor	2007	Strong	Using longitudinal data, found teacher experience had significantly positive effect on student achievement

Table 2 (*cont.*)

Author(s)	Year(s)	Indicator	Findings
Croninger, Rice, Rathbun, & Nisho	2007	Strong	Using longitudinal data, found significant positive effects for years of experience on reading achievement
Kane, Rockoff, & Staiger	2008	Strong	Teacher experience had significantly positive effect on student achievement
Koedell & Betts	2007	Weak	Years of experience only weakly associated with student achievement after finding a small, significant relationship in math but no relationship in reading
Harris & Sass	2008	Weak	Replicating Clotfelter, Ladd, & Vigdor (2006), found half the standard deviation increase and diminishing effect after teachers; first year

A review of the literature on the effect of teacher years of experience over the last few decades is best described as mixed, with empirical studies finding varying levels of correlation between teacher experience and student achievement. The first and most notable examination of this relationship was undertaken by Eric Hanushek, who consistently found no significant relationship between teacher years of experience and student achievement (1986, 1989, 1991, 1997). In his seminal work, Hanushek (1986) analyzed 147 studies available in the literature that examined the relationship between a number of teacher quality indicators—including teacher experience—and student performance. Only 30% of the subset of 109 studies incorporating teacher experience showed a statistically significant result, leading Hanushek to conclude that the indicator was at best a weak indicator of student achievement and, therefore, teacher excellence.

Subsequent studies by Hanushek (1989, 1997) and others inspired by his conclusions spurned a significant debate on the effect of this teacher quality indicator on student achievement. In a longitudinal study of 84,154 9th grade students in Chicago Public Schools, Aaronson, Barrow, and Sanders (2007) found that teacher observable teacher characteristics—

including teacher experience—at best explained 8% of the variation in student achievement results. Similarly, Koedell and Betts (2007) concluded that teacher experience was only weakly associated with student achievement after an analysis of over 16,000 elementary students in San Diego Public Schools yielded a small but significant relationship in math but no relationship in reading.

Further examination of the effect of teacher experience on student achievement has centered on the influence of early-stage teachers on the results of previous studies in an effort to further contextualize Hanushek's early findings. Rivkin, Hanushek, & Kain (2005), in a study of three 4th-7th grader cohorts in Texas totaling nearly 600,000 students, found no relationship between teacher experience and student performance after the third year. Clotfelter, Ladd, and Vigdor (2006), in a study of 3,842 North Carolina 5th grade teachers, found that nearly half of the significant one-tenth of a standard deviation increase shown by students with experienced teachers was explained by effects of initial and second year teachers. In replicating Clotfelter et al. with data from a Florida school district of medium size, Harris and Sass (2008) only found half the standard deviation gains for students with experienced teachers and also concluded that there existed "diminishing increases in teacher productivity" (pg. 19) after teachers' first year. Hanushek and Rivkin (2007), Rockoff (2004), Chingos and Peterson (2011), and Henry, Fortner, and Bastian (2012) report similar findings regarding the diminishing effect of teacher experience on student achievement beyond the first few years, contributing to the critique of the indicator as neither identifying nor incentivizing teacher excellence (Brodsky, DeCesare, & Kramer-Wine, 2010; Edmonds, 2012; Goldhaber, Dearmond, & Deburgomaster, 2011).

While the work of Hanushek ignited a collection of research that showed teacher experience was at best weakly associated with student achievement, others in the field have

mounted a counter-effort that provides evidence that the indicator does impact student achievement, after all. To that end, a number of researchers have completed re-analyses of Hanushek and others' work and raised questions regarding the methodological approaches through which conclusions about teacher experience were drawn. Greenwald, Hedges, and Laine (1996) concluded that teacher experience and student achievement were related, after a re-analysis of Hanushek's (1986) 30-year investigation of teacher quality studies. In addition to leveraging more advanced meta-analysis techniques that did not incorporate "vote counting" (Hedges, Laine, & Greenwald, 1994, p. 6), Greenwald et al (1994, 1996) and others contended that the inclusion of statistically insignificant studies in Hanushek's work significantly influenced his findings and conclusions (Monk, 1989; Zhang, Verstegen, & Kim; 2008).

Building upon movement in the field thereafter towards advanced and alternative meta-analysis techniques, a number of studies in the literature have concluded that teacher experience and student achievement are significantly related. Dewey, Husted, and Kenny (2000), in using meta-analysis to reanalyze Hanushek's (1986) review of teacher quality indicators, also found that teacher experience represented a positive significant effect. In a study of over 5,900 1st – 3rd graders in Tennessee, Nye, Konstantopoulos, and Hedges (2004) found significant positive effects for teacher experience on 2nd grade reading and 3rd grade math achievement. Kane, Rockoff, and Staiger (2008) reached similar conclusions after their study of over 600,000 students in New York City Public Schools showed significant positive effects for teacher experience. Croninger, Rice, Rathbun, and Nishio (2007) reached a similar conclusion after their longitudinal study of 5,167 students found positive effects for reading achievement. Lastly, Clotfelter, Ladd, and Vigdor (2007b), using longitudinal data and building upon their previous research on 3rd – 5th grade students in North Carolina (Clotfelter, Ladd, & Vigor, 2006),

concluded that teacher experience had a significantly positive effect on student achievement. Taken in totality, these counterarguments demonstrate that the debate around this indicator is likely to continue for some time.

This section introduced one of the key teacher quality indicators (i.e. signals), teacher experience, investigated in this study. A review of the relevant literature of teacher experience revealed a significant debate on the effect of this indicator on student achievement, as well as the context therein (e.g. early stage versus experienced teachers). The next section will focus on another signal investigated in this study—teacher educational attainment—which, as will be the case in this study, is often investigated alongside teacher experience in research.

Employee Signal 2: Teacher Educational Attainment

While solidly a part of contemporary compensation plans, salary improvements—and the inherent teacher quality value signaled therein—for master's degrees have been an ongoing, albeit evolving inclusion in the single salary schedule. Inspired from the scientific management movement that began in the 1950s, gaining educational credentials—like master's degrees—was an organizational signal of effectiveness based on the assumption that more education in the field of education would lead to improved instructional practice (Kelley, 1997; Podgursky & Springer, 2007). Master's degree attainment as a quality indicator presented many of the same advantages as teacher experience—it was achievable, unbiased, understandable, and popular with teachers— and skyrocketed in popularity as states increasingly incentivized its procurement or even required it for career status or advanced licensure (Goldhaber & Brewer, 1998; Goldhaber, Darmond, & DeBurgomaster, 2011; Hanushek, 2011; Hanushek & Rivkin, 2006; Springer & Gardner, 2010). Table 3 provides a visual summary of the relevant literature investigating the strength of the relationship between master's degree attainment and student achievement.

Table 3

Summary of Literature Review: Master's Degrees

Author(s)	Year(s)	Indicator	Findings
Hanushek	1986, 1989, 1991, 1997	Weak	Consistently found to be a weak indicator of student achievement, at best
Greenwald, Hedges, & Laine	1996	Strong	Master's degrees and student achievement were related after re-analysis of Hanushek's work using advanced meta-analysis techniques
Goldhaber & Brewer	1997	Weak	Negative, insignificant relationship between master's degrees and student achievement
Dewey, Husted, & Kenny	2000	Strong	Master's degrees and student achievement were related after re-analysis of Hanushek's work using advanced meta-analysis techniques
Aaronson, Barrow, & Sanders	2007	Weak	Master's degrees at best explained 1% of variation in student achievement
Clotfelter, Ladd, & Vigdor	2007	Weak	Found the relationship between master's degrees and student achievement to be occasionally negative and insignificant
Croninger, Rice, Rathbun, & Nisho	2007	Weak	No relationship between master's degrees and reading achievement; Negative, insignificant relationship with mathematics achievement
Koedell & Betts	2007	Weak	Insignificant relationship; master's degree attainment was unable to predict student achievement
Knoeppel & Rinehart	2008	Strong	Master's degree attainment a significantly positive predictor of student achievement
Chingos & Peterson	2011	Weak	No relationship existed between master's degree attainment and student achievement
Kane & Staiger	2012	Weak	Master's degrees produced small, insignificant, and occasionally negative relationship to student achievement in Math and ELA

Similar to teacher experience, criticism of master's degrees as a teacher quality indicator centered around alignment with student achievement (Brodsky, DeCesare, & Kramer-Wine, 2010; Goldhaber, 2002; Harris & Sass, 2008; Omps, 2011-20112; Rivkin, Hanushek, & Kain, 2005; Kane & Staiger, 2012). The coursework and skill development of master's degree programs—and the quality of the programs themselves—have varied significantly and regularly

did not align with strategies for improved instructional practice (Kelley, 1997; Springer & Gardner, 2010). Thus, opponents argue that master's degrees, like teacher experience, fail to differentiate between the abilities of teachers who posses and are compensated for them in the single salary schedule (Brodsky, DeCesare, & Kramer-Wine, 2010; Omps, 2011-20112). However, the main criticism of master's degrees as a teacher quality indicator is that their attainment has not consistently shown to improve student achievement (Chingos & Peterson, 2011; Clotfelter, Ladd, & Vigdor, 2006; Coleman, 1966; Hanushek,1986; Harris & Sass, 2011; Rivkin, Hanushek, & Kain, 2005). A review of the literature on the effect of master's degrees over the last few decades is best described as mixed, with empirical studies finding varying levels of correlation between educational attainment and student achievement.

Hanushek's (1986, 1989, 1991, 1997) early work on teacher quality, first introduced in the teacher experience section, provides the most notable examination of this relationship, and has consistently found no significant relationship between educational attainment and student achievement. In his seminal work, Hanushek (1986) analyzed 147 studies available in the literature that examined the relationship between a number of teacher quality indicators— including master's degree attainment—and student performance. Only 6 of the subset of 109 studies incorporating this indicator showed a significantly positive result, leading Hanushek to conclude that the indicator was a poor indicator of student achievement and, therefore, teacher excellence.

Subsequent studies by Hanushek (1989, 1997) and others found similar conclusions and spurred a debate akin to the aforementioned one on teacher experience on the effect of master's degrees on student achievement. In a longitudinal study of 18,000 students, Goldhaber and Brewer (1997) found a negative, insignificant relationship between master's degree attainment

and student achievement. Clotfelter, Ladd, & Vigor (2007b), in a longitudinal study of $3^{rd} - 5^{th}$ graders in North Carolina, also found the relationship between master's degrees and student achievement to be occasionally negative and insignificant. Croninger, Rice, Rathbun, and Nishio (2007) reached a similar conclusion after their longitudinal study of 5,167 students found no relationship between master's degree attainment and reading achievement and a negative, insignificant relationship with math achievement. Similarly, Koedell and Betts (2007) concluded that master's degree attainment was unable to predict student achievement after an analysis of over 16,000 elementary students in San Diego Public Schools yielded an insignificant relationship for the indicator. In a longitudinal study of 84,154 9th grade students in Chicago Public Schools, Aaronson, Barrow, and Sanders (2007) found that educational attainment at best explained 1% of the variation in and had no relationship to student achievement results. Lastly, Chingos and Peterson (2011), in a study of over 37,000 Florida teachers, concluded that no relationship existed between master's degree and student achievement on the Florida Comprehensive Assessment Test.

Debate on this indicator has extended to university selectivity (i.e. prestige) of the master's degree program, as researchers aimed to determine if the relative quality of degree programs explained the lack of relationship between master's degrees and student achievement. While some studies in the literature have shown positive relationships to that end (Ehrenberg & Brewer, 1994; Summers & Wolfe, 1977), recent research has largely shown a lack of relationship between university quality and effectiveness. To that end, Chingos and Peterson (2011) found that college selectivity yielded an inconclusive relationship in FCAT scores, suggesting that the lack of relationship between student achievement and educational attainment was not influenced by the relative quality of the university attended by teachers. Similarly,

Clotfelter, Ladd, and Vigdor's (2006, 2007a) studies in North Carolina found no effect on student achievement by teachers' college selectivity. Lastly, Boyd, Lankford, Loeb, Rockoff, and Wyckoff (2008) concluded in their study of 4th – 8th grade New York City teachers that there existed only a weak relationship between university quality and teacher effectiveness.

As has happened with teacher experience, Hanushek and others' conclusions that master's degrees are at best weakly associated with student achievement spurred a counter effort in the field that aimed to show that master's degrees do, in fact, impact student achievement. Similar to teacher experience, re-analyses of Hanushek and others' work have raised questions regarding the methodologies employed that resulted in weak or no relationship between master's degrees and student achievement. Most notably, Greenwald, Hedges, and Laine (1996), whose criticism of methodological approaches used by others was presented in the previous section, concluded that master's degrees and student achievement were related after a re-analysis of Hanushek's (1986) 30-year investigation of teacher quality studies. Dewey, Husted, and Kenny (2000), in using advanced meta-analysis to reanalyze Hanushek's (1986) review of teacher quality indicators, also found that master's degrees represented a positive significant effect. Lastly, Knoeppel and Rinehart (2008), in a regression and canonical analysis in 193 Kentucky high schools, found master's degree attainment to be a significantly positive predictor of student achievement. Like teacher teachers of experience, the debate around the impact of master's degrees is far from settled; however, unlike teacher years of experience, fewer studies currently exist that demonstrate the indicator having an impact on student achievement.

This section introduced the second key teacher quality indicator (i.e. signal), master's degree attainment, investigated in this study. A review of the relevant literature of master's degree attainment revealed a somewhat one-sided debate on the effect of this indicator on student

achievement, as well as the context therein (e.g. university quality and/or selectivity). The next section will focus on another signal to be investigated in this study: teacher classroom observations.

Employee Signal 3: NC Teacher Evaluation Rubric

While long a part of the teaching profession, teacher observations rose into the spotlight as a mechanism for teacher accountability and compensation following the *A Nation at Risk* report (NCEE, 1983; Podgursky & Springer, 2007; Hazi & Garman, 1988). This elevation of teacher observations was important since teacher characteristics (e.g. experience, master's degrees) failed to either capture in totality what occurred in the classroom or differentiate between teachers of various skill levels (Goldhaber & Brewer, 1997; Hanushek, 1986). Further, as a teacher quality indicator, teacher observations have served to help qualitatively identify teacher excellence and provide validation for student achievement results (Amrein-Beardsley, 2008; Andrejko, 2004; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Kupermintz, 2003; Mangiante, 2011). Lastly, the state of North Carolina incorporates formal teacher observations as a part of its teacher evaluation process (NC State Board of Education, 2012), so it is an indicator (i.e. signal) investigated in this study.

The salient criticism of this teacher quality indicator is centered on the subjective nature of in-person observations, wherein supervisor bias is afforded the opportunity to affect performance reviews (Hanushek, 1986; Lefkowitz, 2000; Riley, 1985-1986; Varma & Stroh, 2001; Wayne & Ferris, 1990). To that end, teachers given difficult teaching assignments, regardless of ability, are more likely to be reviewed poorly by administrators and assigned blame for negative externalities resulting therein (Amrein-Beardsley, 2008; Kupermintz, 2003). Similarly, the subjective nature of in-person observations opens the door for racial and gender

stereotypes of teachers and students to influence evaluator perceptions and performance reviews. There are also concerns regarding the fidelity with which teacher observations occur, as teachers with more experience tend to receive fewer, less rigorous evaluations (Goldrick, 2002; Hazi & Rucinski, 2009).

These criticisms are grounded in the literature that has shown administrators both receive inadequate training on evaluating teachers and, on average, inflate teacher ratings (Cantrell & Kane, 2013; Kane & Staiger, 2012; Loup, Garland, Ellett, & Ruggutt, 1996; Tucker, 1997). Weisburg, Sexton, Mulhern, & Keeling (2009) chronicled this issue in their *Widget Effect* study, where no more than one percent of the 15,000 teachers studied received an unsatisfactory rating. However, a considerable body of research points to the strength of teacher observations in predicting student performance. In a study of $3^{rd} - 6^{th}$ grade inner-city elementary students, Murnane and Phillips (1981) found that a large portion of student achievement data was explained by targeted instructional strategies by effective teachers. Goldhaber & Brewer (1997) replicated this study with secondary school students and had similar findings. These studies highlighted the potential of teacher observations in predicting student achievement, as effective teacher behaviors could be observed and serve as a proxy for overall teacher quality while mitigating limitations of the single salary schedule (e.g. differentiating between the ability of equally qualified teachers) and student achievement (e.g. validating results).

The potential of these findings is furthered by evidence demonstrating school leaders' ability to identify the most and least effective teachers. In a study of 6th grade students in Los Angeles, Armor et al. (1976) found that principals were able to consistently identify the strongest and weakest teachers. Murnane (1975), in a study of 875 inner city children, also found that principal ratings were strongly correlated with student achievement for effective and ineffective

teachers. A recent study by Rockoff and Speroni (2010) of $3^{rd} - 8^{th}$ grade teachers in New York City found that teachers who received the strongest evaluations during the interview process or early in their career produced stronger student achievement gains. Tyler, Taylor, Kane, and Wooten (2010), in a study of 2,071 teachers in Cincinnati, found that improved ratings in the Teacher Evaluation System (TES)—which included a formal evaluation instrument— corresponded with significant student achievement gains.

The literature also reveals that the strong relationship between teacher observations and student achievement continues to persist in studies involving value-added models (VAMs). Sanders and Horn (1994) found a strong relationship between teacher observations and student Tennessee Value-Added Assessment System (TVAAS) scores. In a study of 202 $2^{nd} - 6^{th}$ grade teachers, Jacob and Lefgren (2005) also found a positive and statistically significant relationship between teacher evaluations and value-added scores. Further, in a study of Tennessee's Project STAR program, Dee and Keys (2004) found that higher student achievement in math was associated with teachers who had been promoted on the basis of strong evaluations. Lastly, in their *Measures of Effective Teaching (MET) Project* reports, Kane and Staiger (2012) and Cantrell and Kane (2013) found teacher observations to be reliable predictors of student achievement when multiple observations were completed by highly trained observers.

This section introduced the third key teacher quality indicator (i.e. signal), teacher observations, investigated in this study via North Carolina Teacher Evaluation Rubric ratings. A review of the relevant literature of teacher observations revealed a considerable amount of support for this indicator as a predictor of student achievement, despite the subjective nature of its construction. The next section will focus on the use of VAMs as a mechanism for determining

teacher quality, as the final signal investigated in this study—EVAAS scores—is constructed on a value-added framework.

Employer Signal 4: EVAAS Rating of "Exceeds Expected Growth"

Due to the subjectivity concerns of teacher observations and the relative lack of student outcome predictability by traditional teacher characteristics, the field of education has actively searched for more objective ways to provide school and teacher accountability by attempting to place more emphasis on student outcomes than teacher inputs (Kelley, 1997; Podgursky & Springer, 2007). As a result, the use of value-added data has increased, as the prospect of measuring students based on growth, rather than proficiency, gained in popularity. Value-added data are used as a mechanism to achieve this objective by calculating the amount of learning a classroom teacher has provided for the student during the school year (Alicias, 2005; Ballou, 2012; Bastian, Henry, & Thompson, 2013). The value added by the teacher, then, is a quantifiable gain, compared to a student's score the previous year, which is attributed to the teacher's instruction (Amrein-Beardsley, 2009; Ballou, 2012; Bastian, Henry, & Thompson, 2013; Costello, Elson, & Schacter, 2008). Most value-added data are generated by complex mixed-method computations that aim to control and adjust for differences in student background and influences (Buddin, 2011; Fuhrman, 2010; McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Sanders & Horn, 1998). The byproducts of this approach are data that describe the impact a teacher has on student growth and performance.

Often the result of multiple years of student scores, value-added data give teachers and school administrators an indication of the effect a teacher had on students meeting or exceeding expected growth (Amrein-Beardsley, Collins, Polasky, & Sloat, 2013; Buddin, 2011; McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Sanders & Horn, 1998). Because value-

added data are generated from comparing a student to his performance from previous year(s), individual students serve as the control in the statistical computation of their growth, as demographic information such as race, ethnicity, and socioeconomic status presumably remain constant each year (Hill, 2011; Sanders, 1998). After first being used to determine farm productivity, value-added data became a fixture in education in the early 1980s when William Sanders adapted this statistical approach for use in schools. Then a professor at the University of Tennessee, Sanders created a model that not only tracked student achievement from year to year, but also used it to project student growth on future standardized tests while controlling for demographic factors (Ballou, Sanders, & Wright, 2004; Stewart, 2006; Wright, Horn, & Sanders, 1997). In 1992, the Tennessee legislature commissioned Sanders' model for statewide use in the Tennessee Educational Improvement Act, which is still used in school accountability today. Since then, a number of states have employed similar value-added models to evaluate their schools and teachers (Amrein-Beardsley, Collins, Polasky, & Sloat, 2013; Stewart, 2006). North Carolina has been a leader in the use of VAMs to identify teacher quality through its state-wide use of VAM product SAS EVAAS (U.S. Department of Education, 2006), and is poised to continue its efforts in this area (Collins & Amrein-Beardsley, 2014; Dewitt, 2014).

One of the main benefits of using value-added data is that it aims to create a level playing field for teacher evaluation while lessening the impact of institutional inequities in the system that impede accurately measuring student and teacher performance. By using a model that calculates student growth each year, while controlling for demographic factors, teachers are not punished for having a classroom with a number of traditionally low-achieving students (Amrein-Beardsley, 2008; Ballou, 2002; Ballou, Sanders, & Wright, 2004). Because value-added data take into account past student performance, teachers are not described as ineffective simply

because their students do not meet proficiency standards. Rather, their effectiveness is determined by their performance relative to previous year(s) performance (Amrein-Beardsley, Collins, Polasky, & Sloat, 2013; Buddin, 201; David, 2010). For these reasons, value-added models have become an attractive indicator of teacher quality, as proponents argue that VAMs allow for a re-focus on expected, realistic progress and equitable teacher evaluations (Alicias, 2005; David, 2010; Misco, 2008).

Criticism of value-added data generally fall into three major categories: sampling, validity and reliability, and year-over-year output variability (Amrein-Beardsley, 2009; Amrein-Beardsley, Collins, Polasky, & Sloat, 2013; David, 2010). With respect to sampling, opponents argue that VAMs fail to control for the non-random nature of teacher and student assignment (Amrein-Beardsley, 2009; Chingos & Peterson, 2011; David, 2010). When coupled with the tendency for stronger students to be assigned to teachers deemed to be more effective, the non-random nature of student assignment has the potential to skew value-added results and undermine efforts to capture the actual value added by teachers (Amrein-Beardsley, 2009; Chingos & Peterson, 2011; Clotfelter, Ladd, & Vigor, 2006; Lockwood, McCaffrey, and Sass, 2008; Rothstein, 2010). Another sampling concern relates to the size of the samples used in determining value-added teacher ratings, as value-added ratings stemming from smaller sample sizes have the potential for greater variability despite comparable teacher effectiveness (Amrein-Beardsley, 2009, Kane & Staiger, 2001; Lockwood, McCaffrey, & Sass, 2008). For this reason, many researchers recommend that three or more years of data be included in value-added calculations (Brophy, 1973; Cody, McFarland, Moore, & Preston, 2010, Harris, 2011).

Critics also point to validity and reliability as areas of concern with VAMs and the standardized tests connected to them. In addition to modeling and statistical error inherent in

standardized testing, opponents argue that it is difficult to compare scores from consecutive years

because most tests do not measure the same content, which is only confounded by serious doubts

that individual tests actually measure the amount of content learned in a course in one year or

even comprehensively represent the curriculum (Amrein-Beardsley, 2009; David, 2010; Misco,

2008; Rothman, 2010). Due to concerns about flaws inherent in the assessments used, there have

been significant challenges to the validity of the value-added data generated from them.

Proponents and developers of VAMs, however, have pointed to studies where VAM

scores were strongly correlated to other measures of teacher quality (e.g. teacher observations) as

evidence of the validity and reliability of the models used (Chingos & Peterson, 2011; Harris &

Sass, 2008, Jacob & Lefgren, 2008; Kane & Staiger, 2008; Nye, Konstantopoulos, & Hedges,

2004). In a study of 3,000 teachers from six urban districts, findings from the *Measures of*

Effective Teaching (MET) Project provided additional evidence for the validity of VAMs after

finding strong relationships between high teacher value-added scores and strong teacher

observations with non-randomized (Bill & Melinda Gates Foundation, 2010) and randomized

(Kane & Staiger, 2012) teacher assignments. Students of high value-added rating teachers also

tended to perform higher on supplemental tests with higher cognitive complexity, providing

additional evidence for high value-added scores as a measure of effective teaching (Cantrell &

Kane, 2013). These findings aside, the body of research on VAMs has largely presented

evidence that fuels validity and reliability concerns, rather than confidence (Amrein-Beardsley &

Collins, 2012; Au, 2010; Koedel & Betts, 2007; Papay, 2010).

Lastly, critics point to year-over-year variability in value added scores as a sincere

limitation of the models used. Recent studies have shown that many teachers deemed effective

one year were not identified as such the following year (David, 2010; Harris, 2010; Koedel &

37

Betts, 2007). Other studies have concluded that consistent trends in VAM ratings from year to year can be found, but only if students are randomly assigned to teachers (Fuhrman, 2010; Kane & Staiger, 2008; Nye, Konstantopoulos, & Hedges, 2004; Rothman, 2010; Viadero, 2008). These findings cause concern because should value-added results be a measure of teacher quality, the lack of consistency, or control over causes of inconsistency, could lead to teachers being incorrectly identified as low performing. Furthermore, both student and teacher performance can fluctuate considerably throughout the school year and beyond, making accurate evaluation difficult and eroding the core assumption—stability of student and teacher characteristics— underlying the use of VAMs for determining teacher value added ratings (Edmonds, 2012; Linn & Huag, 2002; Kane & Staiger, 2002). However, while year-over-year fluctuation in value add scores is a reality, annual reports from the *MET* study have shown that incorporating value added ratings increases the predictive power of identifying teachers who yield student achievement gains with subsequent cohorts of students (Bill & Melinda Gates, 2009; Kane & Staiger, 2012; Cantrell & Kane, 2013).

By incorporating value-added measurements, districts are attempting to use the most objective and sophisticated tool they have as a way to identify teacher excellence (Edmonds, 2012; Amrein-Bearsley, 2008). In North Carolina, the VAM movement has manifested into the implementation of SAS EVAAS (Collins & Amrein-Beardsley, 2014; Dewitt, 2014; U.S. Department of Education, 2006). Touted as the most comprehensive and reliable model available by proponents, EVAAS purportedly corrects for many of the criticisms, except for test quality (Mangiante, 2011; Ravitch, 2010), previously mentioned in addition to controlling for student background and missing data (Amrein-Beardsley, 2008; Collins, 2014; SAS, 2012). However, a dearth of evidence exists to support the superiority claims of EVAAS, largely because the model

is proprietary and not publicly available for critique in its totality (Amrein-Beardsley, 2008; Collins, 2014). Additionally, in a study of teachers in Houston whose contracts were not renewed in-part based on EVAAS ratings, Amrein-Beardsley and Collins (2012) concluded that validity and reliability issues prevalent in traditional VAMs also appeared in teacher EVAAS ratings. A follow up study by Collins (2014) of Houston's use of EVAAS revealed similar validity and reliability findings, as no evidence existed to back up claims that the model could control for "extraneous variables such as home life, health, behavior, motivation, etc. on student achievement" (p.20). Thus, while some studies have provided evidence of the long-term impact of high value-added teachers (Chetty, Friedman, & Rockoff, 2011; Chingos & West, 2012; Rockoff & Speroni, 2011), it has become increasingly difficult to discern the amount of value added by the teachers (Ishii & Rivkin, 2009; Kennedy, 2010; Linn, 2008; Papay, 2010).

This section focused on the use of VAMs as a mechanism for determining teacher quality, as the final signal to be investigated in this study, EVAAS scores, is constructed on a value-added framework. A review of the relevant literature revealed considerable debate on the accuracy of value-added scores and the teacher quality ratings created from them. Despite being the most objective measure of student performance, critics present sampling, validity and reliability, and variability concerns that thus far have not been sufficiently addressed in the literature. The next section will present a summary of the literature reviewed in Chapter 2.

Summary

This chapter provided a review of the conceptual lens used, and the relevant literature on the teacher quality indicators (i.e. signals) investigated, in this study. Spence's (1974) Signal Theory has a long history of application in other disciplines and was chosen by the researcher for this study due to its capacity to reduce informational asymmetries between teachers and

educational leaders on underlying teacher quality. A review of relevant literature on the signals

most often used by teachers and educational leaders—years of experience, master's degree

attainment, North Carolina Teacher Evaluation Rubric ratings, and EVAAS scores—was also

presented, and revealed considerable debate on the efficacy of each in serving as a proxy for

teacher quality. As such, this study adds to the literature of each signal, as well as informs

educational leaders in North Carolina interested in creating differentiated teacher compensation

plans using the indicators that most accurately predict student growth.

CHAPTER III: DATA & METHODOLOGY

This chapter presents the data and methods used to complete the study, and is separated into sections that detail the purpose of the study, research questions, population and sample, data procedures, validity and reliability, research design framework, research design, and framework used for analysis.

Purpose of the Study

Given North Carolina's ongoing need to identify quality teachers in a marketplace wherein candidate characteristics are consistently and finitely signaled, the purpose of this study was to investigate the relationship between those signals of teacher quality in North Carolina— number of years of experience, master's degrees, North Carolina Teacher Evaluation Rubric rating—and student performance, measured by EVAAS scores. Specifically, the goal was to determine signal symmetries and asymmetries between the state's outcome measure and the available teacher quality indicators serving as predictors to that end. That is, in what way are the indicators leveraged by teachers and allowed by the state to preemptively demonstrate teacher quality related to the state's established criteria for determining teacher excellence?

Research Questions

1. In what way, if any, do North Carolina's teacher quality indicators significantly predict student growth, as measured by EVAAS?

2. What is the profile of the North Carolina teacher whose students exceed expected growth?

The sample used in this study was drawn from the population of full-time teachers in North Carolina's public schools from the 2010-2011 through 2013-2014 school years with observations for teacher years of experience, master's degrees, North Carolina Teacher Evaluation Rubric rating, at least one EVAAS score in a given year, and teacher and school demographic data, yielding a total of 101,349 observations. Descriptive statistics of the sample are presented later in this section and a full list of variables is found in Appendix A. The next paragraph describes how data were collected and compiled.

Data were collected from secondary data sets housed by and with the permission of the Education Policy Initiative at Carolina (EPIC), as well as publicly available data from the Elementary and Secondary Information System (ELSI) at the National Center for Education Statistics (NCES) and the North Carolina Department of Public Instruction (NCDPI). Teacher and school variables were matched on existing generic, randomly assigned teacher identification numbers provided by EPIC, standardized school and district local education agency (LEA) codes, and year of observation. Leveraging secondary data sources, which consolidate data used by education agencies in determining teacher quality and education policy, was the most practical method for collecting and leveraging a representative sample of teachers for this study. In order to ensure the protection of personally identifiable information, the researcher followed security protocols described in a Memorandum of Agreement (MOA) with EPIC and the North Carolina Department of Public Instruction (NCDPI) found in Appendix B. The next paragraph describes properties of the data set.

The total number of complete teacher observations, as well as the number of observations for individual variables, fluctuated each year. Most variables saw an increase in observations

over time, which can be attributed to an increase in the number of teachers whose students took

EVAAS-generating state assessments as well as a more wide-ranging implementation of the

North Carolina Teacher Evaluation Rubric. The resulting panel (longitudinal) data set was

therefore considered unbalanced; actions by the researcher to control for the unbalanced nature

of the data set, as well as other threats to interpretation, are found later in this chapter. The

paragraphs that follow outline the key teacher quality variables investigated, as well as additional

teacher and school variables used in in modeling.

Teacher Years of Experience

Teacher experience has long been used by researchers and policymakers as a proxy for

teacher quality though its inclusion in state single salary schedules (Kelley, 1997; Omps, 2011-

2012, Springer & Gardner, 2010) and is still a metric that informs quality, compensation, and

employment (e.g. tenure) designations in North Carolina (NC Department of Public Instruction,

2015; NC State Board of Education, 2012). This variable, measured in years, was represented in

the data as the number of years a teacher had taught in the public-school system in North

Carolina in a given school year. The average teacher years of experience in each year of was

consistently between 11 and 12 years with a standard deviation between 8.8 and 9.2. Table 4

provides a visual summary of this variable for each year in the panel set.

Table 4

Teacher Years of Experience by Year

	n	Mean	SD	Min	Max
2011	22,879	11.17	8.99	0	50
2012	28,071	11.69	9.15	0	55
2013	38,348	11.20	8.91	0	52
2014	59,982	11.07	8.85	0	53
Overall	149,280	11.24	8.95	0	55

Teacher Master's Degrees

Based on the notion that more education would lead to improved instructional practice (Kelley, 1997; Podgursky & Springer, 2007), states have increasingly incentivized master's degree procurement or even required it for career status or advanced licensure (Goldhaber & Brewer, 1998; Goldhaber, Darmond, & DeBurgomaster, 2011; Hanushek, 2011; Hanushek & Rivkin, 2006; Springer & Gardner, 2010). Teacher master's degrees were represented as either 0, indicating the teacher did not have an advanced degree, or 1, indicating the teacher did have an advanced degree, in a given school year. The researcher decided to consolidate all advanced degrees together in order to determine if advanced degrees more broadly impacted student performance rather than narrowly focusing on master's degree attainment alone.

The mean value for teacher master's degrees in each year investigated was between .33 and .39 which, given the binary nature of the variable indicates that teachers with master's degrees made up no more than 39% of teachers in a given year. Table 5 provides a visual summary of this variable for each year in the panel set.

Table 5

Teacher Master's Degrees by Year

	n	Mean	SD	Min	Max
2011	23,631	.34	.47	0	1
2012	28,800	.38	.48	0	1
2013	39,486	.39	.49	0	1
2014	61,477	.37	.48	0	1
Overall	153,394	.37	.48	0	1

Teacher Observations: North Carolina Teacher Evaluation Rubric Ratings

The state of North Carolina incorporates formal teacher observations as a part of its teacher evaluation process (NC State Board of Education, 2012), which follows a trend in education to use this variable to help identify teacher quality and provide validation for student achievement results (Amrein-Beardsley, 2008; Andrejko, 2004; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Kupermintz, 2003; Mangiante, 2011). North Carolina Teacher Evaluation Rubric ratings were represented in the data as a composite of the ratings teachers received each year on the five standards described in the North Carolina Teacher Evaluation Rubric (NC State Board of Education, 2012). The standards and elements are shown in the full rubric found in Appendix D. The paragraphs that follow describe the process by which teachers receive these ratings and how the variable was created.

At the end of each school year, each teacher's principal provides an overall rating of Not Demonstrated, Developing, Proficient, Accomplished, or Distinguished for each standard in the rubric based on progress on the elements and artifacts presented by the teacher. These standard

ratings correspond to values of 0-4 in the panel set, respectively. For the purposes of having a single variable representing teacher evaluations each year, the researcher used the median score of standard ratings for each teacher in each year to create a composite score used in analysis. Leveraging a median to create a composite score has been done in prior investigations by EPIC, as evidenced by this composite score being a pre-existing variable in a portion of data sets shared with the researcher where medians where used to calculate composites

The mean composite score was between 3.6 and 3.7, yielding an average rating of "Accomplished" in each year. Table 6 provides a visual summary of this variable for each year in the panel set.

Table 6

Teacher Evaluation Rubric Ratings by Year

	n	Mean	SD	Min	Max
2011	11,419	3.56	.67	1	5
2012	15,891	3.60	.68	1	5
2013	38,324	3.73	.68	1	5
2014	55,148	3.70	.68	1	5
Overall	120,782	3.68	.48	1	5

Teacher EVAAS Scores

North Carolina has been a leader in the use of VAMs to identify teacher quality through its state-wide use of VAM product SAS EVAAS (U.S. Department of Education, 2006) and is a component of the North Carolina Teacher Evaluation Process (NC State Board of Education (2012). Teacher EVAAS scores were indexed in the data set so that teacher scores from various grade levels and subjects could be used for analysis in answering Research Question 1, which aimed to determine the relationship between the key teacher quality variables described in prior sections and student performance. Table 7 provides a visual summary of the resulting variable, *Index*, for each year in the panel set.

The mean indexed EVAAS score was between .09 and .14, which is expected given that EVAAS scores are essentially z-scores centered at zero. An average standard deviation of 2.52 to 2.9 is higher than expected, however, as indexed scores greater than or equal to 2.0 represent exceeding expected growth. Mean scores also increased over time while standard errors mostly decreased, which implies that while teacher performance improved the difference between teachers' performance lessened.

Table 7

Indexed Teacher EVAAS Scores by Year

	n	Mean	SD	Min	Max
2011	23,705	.09	2.90	-14.61	19.99
2012	29,254	.06	2.52	-17.83	19.20
2013	39,839	.11	2.88	-33.59	18.86
2014	61,803	.14	2.56	-23.13	23.81
Overall	154,601	.11	2.69	-33.59	23.81

Indexed Teacher EVAAS Scores Exceeding Expected Growth

Indexed Teacher EVAAS scores were used to create a new variable that categorized teachers as either exceeding expected growth or not exceeding expected growth. This designation was determined by indexed teacher EVAAS scores greater than or equal to 2.0, representing two standard deviations above the mean EVAAS score for a single state test. The resulting variable, *IndexExceeds*, was therefore assigned a value of 0 for indexed EVAAS scores less than 2.0, and assigned a value of 1 for scores greater than or equal to 2.0. This dependent variable was essential in answering Research Question 2, which aimed to determine the profile of teachers whose students exceeded expected growth.

The percentage of teachers who exceeded expected growth was stable, with mean values between .17 and .19 for each year investigated, even as the number of teachers studied each year fluctuated. This corresponds to be tween 17% and 19% of teachers per year exceeding expected growth. Table 8 provides a summary of *IndexExceeds* for each year in the panel set.

Table 8

Indexed Teacher EVAAS Scores Exceeding Expected Growth by Year

	n	Mean	SD	Min	Max
2011	23,705	.19	.39	0	1
2012	29,254	.17	.37	0	1
2013	39,839	.19	.39	0	1
2014	61,803	.19	.39	0	1
Overall	154,601	.18	.39	0	1

The following sections describe demographic teacher and school variables that accompanied

teacher quality observations in the data set. These variables included teacher gender and racial

minority status, as well as school Title 1 status, non-White rate, enrollment, and urbanicity.

Teacher Gender

The variable for gender was represented in the panel as either 0, indicating the teacher

was male, or 1, indicating the teacher was female. All gender-related data shared with the

researcher described teachers as either male, female, or had no gender listed (i.e. missing). This

variable was an important inclusion in the model for Research Question 2, as the researcher

aimed to determine the impact of teacher gender on the relative probability of a teacher

exceeding expected growth.

The percentage of female and male teachers was fairly stable for the years investigated,

with the percentage of female teachers between 77% and 83%, even as the number of teachers

studied each year fluctuated. Table 9 provides a summary of teacher gender for each year in the

panel set.

Table 9

Teacher Gender by Year

	n	Mean	SD	Min	Max
2011	23,571	.77	.42	0	1
2012	29,097	.82	.39	0	1
2013	39,648	.77	.42	0	1
2014	61,457	.83	.37	0	1
Overall	153,773	.80	.40	0	1

Teacher Racial Minority Status

The variable for teacher racial minority status was represented as either 0, indicating the teacher was not a racial or ethnic minority, or 1, indicating the teacher was a racial or ethnic minority. The literature discussed in Chapter 2 largely avoided the intersection of teacher race and ethnicity on student performance; however, the researcher wanted to ensure that the model(s) used accounted for its impact on the relative probability of a teacher exceeding expected growth. Further, this variable was constructed as binary to account for the impact of White privilege in education institutions, policies, purported pedagogical best practices, and perceptions of teacher quality on teacher evaluation and performance (Delpit, 1995; Epstein, 2005; Goldhaber & Hansen, 2010; Ladson-Billings, 2006). However, this approach limits the generalizability of the findings presented, as the researcher is not able to identify the impact of a teacher belonging to various racial and ethnic minority groups on performance indicators. Subsequent research could therefore investigate the stability of the findings presented in this study across different teacher racial minority subgroups.

The percentage of racial minority teachers was stable, with percentages between 18% and 19%, for each year investigated even as the number of teachers studied each year fluctuated. Table 10 provides a summary of teacher racial minority status for each year in the panel set.

Table 10

Teacher Racial Minority Status by Year

	n	Mean	SD	Min	Max
2011	23,694	.18	.38	0	1
2012	29,248	.19	.39	0	1
2013	39,832	.18	.39	0	1
2014	61,763	.18	.38	0	1
Overall	154,537	.18	.39	0	1

The following sections describe school variables that accompanied teacher quality observations in the data set and include school Title 1 status, non-White rate, enrollment, and urbanicity. These school variables had more 2011 observations than teacher variable observations for that year in the data set, but would mirror teacher variable observation totals in subsequent years. This spike in 2011 observations is best described as a clustering of school variable observations that were connected to missing values for teacher quality and demographic variables.

School Title 1 Status

Serving as a proxy for school poverty, school Title 1 status was represented in the panel as either 0, indicating the school was not Title 1 eligible, or 1, indicating the school was Title 1 eligible. Schools are considered Title 1 eligible in North Carolina if at least 40% of their students come from poor families as evidenced by qualifying for free and reduced priced lunch (North Carolina Department of Public Instruction, 2017). This variable was an important inclusion in the model for Research Question 1, as the researcher aimed to control for school poverty in determining the relationship between teacher quality variables and indexed EVAAS scores, and

Research Question 2, as the researcher aimed to determine the impact of school poverty on the relative probability of a teacher exceeding expected growth.

The percentage of Title 1 eligible schools was somewhat stable, with percentages between 81% and 84%, for the years investigated. Table 11 provides a summary of teacher observations broken down by Title 1 status for each year in the panel set.

Table 11

School Title 1 Status by Year

	n	Mean	SD	Min	Max
2011	89,585	.83	.37	0	1
2012	29,254	.81	.39	0	1
2013	39,839	.81	.39	0	1
2014	61, 803	.84	.36	0	1
Overall	220,481	.83	.38	0	1

School Non-White Rate

Serving as a proxy for school demographics, school Non-White rate was represented in the panel set as the percentage of non-White students in a school. Similar to teacher minority race, the teacher quality literature discussed in Chapter 2 largely avoided the impact of school racial demographics on student and teacher performance; however, the researcher wanted to ensure that the model(s) used accounted for its impact on the relative probability of a teacher exceeding expected growth. This variable was constructed as binary to account for the byproducts of White privilege and racial bias in traditional curriculum, tracking and placement, and standardized assessments (Archbald & Farley-Ripple; Battey, 2013; Brown, 2013;

Cornbleth, 2006; Santelices & Wilson, 2010; Steele, 2003). Subsequent research could therefore investigate the stability of the findings presented in this study across different student racial minority subgroups. However, this approach limits the ability of the researcher to identify the impact of a student belonging to various racial and ethnic minority groups on teacher performance indicators. The average school Non-White rate was relatively stable, with mean values .45 and .52, in the years investigated. Table 12 provides a visual summary of this variable for each year in the panel set.

Table 12

School Non-White Rate by Year

	n	Mean	SD.	Min	Max
2011	88,995	.52	.23	0	1
2012	28,847	.49	.25	0	1
2013	37,022	.47	.25	0	1
2014	57,494	.48	.24	0	1
Overall	212,358	.50	.24	0	1

School Enrollment

The school enrollment variable was represented as the total number of students at the school. The average school enrollment saw increases year over year, resulting in more than a 220-student enrollment gain in the years investigated. This variable was an important inclusion in the model for Research Question 1, as the researcher aimed to control for school size in determining the relationship between teacher quality variables and indexed EVAAS scores, and

Research Question 2, as the researcher aimed to determine the impact of school size on the relative probability of a teacher exceeding expected growth.

School enrollment also served as a proxy for school type in this study, as moving from elementary, to middle, to high school indicates an increase in school size, on average. However, the effectiveness of this proxy variable may be limited, as the average enrollment for elementary, middle, and high schools vary greatly between urban, suburban, town, and rural settings. Table 13 provides a visual summary of this variable for each year in the panel set.

Table 13

School Enrollment by Year

	n	Mean	SD	Min	Max
2011	89,004	688.24	417.00	0	2686
2012	28,847	851.27	430.82	19	2766
2013	37,035	895.12	495.71	0	2776
2014	57,494	838.81	439.51	16	2775
Overall	212,380	787.22	447.92	0	2776

School Urbanicity

The urbanicity variable was assigned to schools using the NCES Urban-Centric Locale categories related to school zip codes. The codes were then consolidated into four categories, with city schools represented with a 1, suburban schools represented with a 2, town schools represented with a 3, and rural schools represented with a 4. These designations created a spectrum for school urbanicity in that moving from lower to higher values indicated a move from larger, more urban settings to smaller, more rural settings (ArcGIS, 2016).

While the number of observations fluctuated year over year, the mean value for urbancity remained fairly constant between 2.30 and 3.0, indicating that the average school was in a small suburban area or large town. This variable was an important inclusion in the model for Research Question 2, as the researcher aimed to determine the impact of school urbanicity and rurality on the relative probability of a teacher exceeding expected growth. Table 14 provides a visual summary of this variable for each year in the panel set.

Table 14

School Urbanicity by Year

	n	Mean	SD	Min	Max
2011	88,705	2.30	1.35	1	4
2012	28,926	3.00	1.25	1	4
2013	37,061	2.64	1.25	1	4
2014	57,531	2.78	1.25	1	4
Overall	212,380	2.59	1.32	1	4

Data Procedures

The researcher chose to create a panel data set to study multiple teacher quality and teacher and school demographic variables over time. In merging data sets, the researcher dropped observations that

- Occurred outside the 2010-2011 through 2013-2014 school year study period; and

- Related to charter schools, as the focus of this study was limited to traditional public schools.

Each teacher used in analysis had at least one state assessment that yielded an EVAAS score, representing the calculated value-added gain by the teacher for a section of students on the state assessment. Each teacher EVAAS score was accompanied by a standard error, which was used by the researcher to create a simple index using the formula

$$Index = \frac{EVAAS\ Score}{EVAAS\ SE} \qquad (1)$$

where EVAAS Score is the pre-calculated value-added gain and EVAAS SE is the pre-calculated standard error. In cases where teachers had multiple EVAAS scores in a given school year, the researcher created a composite index based on the technical guidelines provided by the North Carolina Department of Public Instruction (2016). A summary of these procedures is found in Appendix C.

In order to confirm accuracy of these guidelines, the researcher compared calculated composite index scores with observations in the data sets provided by EPIC that included a composite index score. The researcher was able to replicate the index composite scores provided by EPIC to within 0.01. Given the consistency and closeness that resulted, the researcher created a new "Index" variable for all teachers in the panel set to be used in modeling and analysis. Consistent with the guidelines used, composite index values for teachers with only one test yielding an EVAAS score were identical to the simple index using the formula described above.

Validity and Reliability

While the researcher used no instrument in the completion of this study, it is important to address validity and reliability with respect to the secondary data collected and used. A portion of the data used in this study was collected from secondary data sets housed by and with the permission of the Education Policy Initiative at Carolina (EPIC). The Education Policy Initiative at Carolina (EPIC) is a policy group operating under the Public Policy Program at the University

of North Carolina at Chapel Hill. EPIC aims to "conduct policy-relevant research and program evaluation to inform local, state, and federal programs and policies focused on improving effective teaching and student achievement in North Carolina and the nation" (EPIC, 2015). For the last several years, EPIC has assisted the state of North Carolina in completing research projects pursuant to the state's compliance with the federal Race to the Top (RttT) grant. To that end, EPIC has housed secondary data sets containing teacher quality and demographic data on full time public school teachers in North Carolina. Other secondary sets used in this study come from publicly available data housed by the Elementary and Secondary Information System (ELSI) at the National Center for Education Statistics (NCES) and the North Carolina Department of Public Instruction (NCDPI).

Thus, the data used by the researcher can rationally be considered valid and reliable, as EPIC, ELSI/NCES, and NCDPI have already certified the data as such pursuant to their own research priorities. The designation of the data as valid and reliable therefore also extends to individual data points in the set, which mitigate concerns about validity and reliability of data points that leverage standardized testing (i.e. EVAAS scores) and subjective observations (i.e. North Carolina Teacher Evaluation Rubric ratings).

Research Design Framework

The research questions investigated in this study involve investigating the relationship between independent and dependent variables such that the former predict the latter. The statistical tool most appropriate to answer and analyze these questions therefore involves a quantitative, multiple regression approach. According to Creswell (2012), "multiple regression is a statistical procedure for examining the combined relationship of multiple independent variables with a single dependent variable" (p. 350). While regression has long been a part of the field that

explored the relationship between educational inputs and student achievement, Goldhaber and Brewer (1997) championed methodological improvements that addressed unobservable biases found in educational research. To that end, their seminal study incorporated large, longitudinal (i.e. panel) data sets and "fixed-effects" which controlled for unobservable teacher, school, and classroom characteristics (p. 510). While not correlated with observable variables, these fixed effects created a more robust regression model that helped explain variation in student achievement, yielding the conclusion that educational attainment was significant in math classrooms where the teacher had a mathematics bachelor's or master's degree.

The trend of integrating fixed effects and multi-year data sets into traditional regression models investigating the relationship between classroom characteristics and student achievement continued with Hanushek, Kain, & Rivkin (1998), who concluded that master's degrees had no effect on student achievement and years of experience had a decreasingly significant effect over time. Similar approaches to regression modeling by Jacques & Brorsen (2002), Croninger, Rice, Rathbun, and Nishio (2007), Clotfelter, Ladd, and Ligdor (2007), also presented in Chapter 2, have contributed to the literature on the relationship between teacher quality indicators and student achievement.

Panel data and fixed effects integration in regression modeling continued even as value-added modeling became more commonplace. Goldhaber and Anthony (2007), in a study of more than 770,000 North Carolina students, used this methodological approach to conclude that teacher certification through the National Board for Professional Teaching Standards (NBPTS) had little effect on elementary student achievement. Shortly thereafter, Goldhaber (2007) would again leverage panel data and fixed effects in a value-added regression model that included over 1 million teacher-student observations over a 10-year period to conclude that a small, positive

relationship existed between scores on teacher licensure exams and subsequent student achievement.

Value-added modeling in teacher quality research, and the field of education more broadly, has continued to be a controversial issue since. Leveraging modeling techniques developed over the previous decade that incorporated panel data and teacher fixed effects in value-added regression models, Rothstein (2009) concluded that factors leading to student-teacher assignments could contribute greatly to bias inherent in value-added models. Rothstein (2010) extended these findings by creating and testing falsification tests for value-added model integration in regression analyses examining teacher inputs on student achievement, concluding that controlling for unobservable school-level characteristics is also needed in models that include value-added modeling data. Of note, the value-added model used in these studies by Rothstein was the Tennessee Value Added Assessment System (TVAAS), which forms the basis of the EVAAS model investigated in this study (SAS, 2016).

In summary, the literature supports the integration of panel data and fixed effects to control for unobservable characteristics at the teacher, school, and district level in regression analyses that investigate the relationship between teacher characteristics and student achievement in the value-added modeling era. To that end, the next section describes the research design used by the researcher that integrates the research base and recommendations presented in this section pursuant to robust regression models designed to answer Research Question 1 and Research Question 2.

Research Design

The researcher used a regression model and panel data set of observations from the 2010-2011 through 2013-2014 school years to investigate teacher observations over time, where

individual teachers were represented once, multiple times, or in every year studied. Creating a

panel data set for use with the regression model was therefore critical in order to ensure that

individuals with multiple observations over time did not shrink standard errors, thereby

preserving the fidelity of statistical significance, and to mitigate the impact of teachers with

incomplete data. Regression models commonly used in teacher quality research (e.g. Chingos &

Peterson, 2011; Chingos & West, 2011; Hanushek, 2011; Hanusheck & Rivkin, 2010) vary, but

generally formulate as

$$A = \beta_1 X + \beta_2 C + \beta_3 S + \beta_4 W + \pi + \varepsilon \qquad (2)$$

where A represents student achievement; X is a vector representing student characteristics; C is a

vector representing classroom characteristics; S is a vector representing school characteristics; W

is a vector representing teacher characteristics; π represents fixed effects; and ε represents a

standard error term.

For Research Question 1, the researcher wanted to know if the variables of teacher years

of experience, master's degrees, and North Carolina Teacher Evaluation Rubric rating, whether

combined or individually, significantly predicted EVAAS scores. To support an accurate

regression, EVAAS scores were indexed so that scores from teachers in different grade levels,

subject areas, and years, could be analyzed together. For teachers with multiple indexed EVAAS

scores in a single year, the researcher combined scores into a single index composite score in

accordance with the Technical Documentation for EVAAS Analyses (NC Department of Public

Instruction, 2016). A summary of creating composite index EVAAS scores is found in Appendix

C. To that end, the baseline model was re-written as

$$Index = \beta_1 X + \beta_2 C + \beta_3 S + fe_t + \varepsilon \qquad (3)$$

where *Index* represents indexed teacher EVAAS scores; *X* represents teacher years of experience; *C* represents master's degrees; *S* represents North Carolina Teacher Evaluation Rubric rating; fe_t is a teacher fixed effect term controlling for unobservable and/or demographic teacher characteristics (e.g. race, gender); and ε represents a standard error term. The model was further tested and derived yielding a final model of

$$Index = \beta_1 T + \beta_2 S + fe_d + fe_t + e \qquad (4)$$

where *Index* represents indexed teacher EVAAS scores; *T* represents a vector of teacher quality variables; *S* represents a vector of school characteristics; fe_d represents district fixed effects; fe_t represents teacher fixed effects; and ε represents a standard error term. In this model, *T* includes variables of teacher years of experience, master's degrees, and North Carolina Teacher Evaluation Rubric rating. *S* includes variables for school Title 1 status, percentage of non-white students, and enrollment and is integrated into the model to control for the impact of school-level characteristics on indexed EVAAS scores. The fe_d term uses district LEA identifiers in order to control for district policies and characteristics between school districts that impact teacher quality and student achievement. The fe_t term controls for time-invariant unobservable and/or demographic teacher characteristics (e.g. race, gender), while ε represents a standard error term.

For Research Question 2, the researcher wanted to determine the impact of all teacher variables available, including demographic characteristics, on the relative probability of a teacher having indexed EVAAS scores that exceeded expected growth. To determine which variables to include in the model, the researcher first created a new variable, *IndexExceeds,* that categorized teachers as either exceeding expected growth (value = 1) or not exceeding expected growth (value = 0). This designation was determined by *Index* scores greater than or equal to 2.0, representing two standard deviations above the mean. The researcher then ran two-variable t-

tests with *IndexExceeds* and various teacher quality and teacher and school demographic variables. Teacher variables yielding a statistically significant result at $a < .05$ were included in the subsequent regression model used to investigate Research Question 2.

The model used to answer Research Question 2 differed from Research Question 1 in that the categorical, binary nature of the *IndexExceeds* dependent variable (i.e. teachers either exceeded expected growth or they did not) necessitated a shift to logistic regression. This is a shift away from the traditional teacher quality research presented in Chapter II that investigates the impact of teacher characteristics on student achievement, as most leverage a production functions conceptual lens that looks at incremental changes to student performance as teacher inputs vary. To that end, the preponderance of research in this arena uses a continuous variable for student achievement in order to generate conclusions. However, logistic regression is the most logical model and statistical test to use for Research Question 2 given both the binary nature of the variable and the researcher's interest in determining factors that increase a teacher's probability of exceeding expected growth.

While teacher quality research that investigates teacher impact on student achievement may not often leverage logistic regression, others in the teacher quality arena have used this approach—in particular, with respect to teacher retention. In a study of grades 3-8 teachers in Florida, Ingle (2009) used a logistic regression model that included teacher and school characteristics, as well as value-added scores, to determine the relative probability of teacher attrition. Here, the teacher attrition dependent variable was binary (i.e. 0 = teacher stayed at school; 1 = teacher left school) and Ingle concluded that teachers with higher value-added scores were less likely to leave their school the following year. Vagi, Pivovarova, and Barnard (2017) similarly used a logistic regression model to investigate teacher attrition, but instead used a

composite teacher evaluation instrument score as the teacher quality variable to determine

relative probability and concluded that pre-service teachers with higher ratings on the instrument

were more likely to remain in the profession. For these studies, fixed effects were used for

unobservable teacher and school characteristics and to control for time in the underlying panel

data sets leveraged for analysis. Others in the field of teacher quality research related to teacher

retention have led further development using these techniques (e.g. Boyd, Grossman, Lankford,

Loeb, & Wyckoff, 2011; Krieg, 2006).

In summary, a body of research exists wherein teacher quality variables, value-added-

data, and teacher and school characteristics were combined with various fixed effects in logistic

regression analyses. Precedent also exists for leveraging panel data in these models provided that

controlling for time is included in the model. Lastly, the baseline logistic regression model

varies, but generally formulates as

$$log \frac{p(x)}{1-P(x)} = \beta_0 + x \cdot \beta \qquad (5)$$

where $log \frac{p(x)}{1-P(x)}$ represents the conditional probability of the binary outcome investigated and

$\beta_0 + x \cdot \beta$ represents the line of demarcation between the two variables; that is, a probability of

zero.

Applying this baseline logistic regression model to Research Question 2, where the

researcher aims to determine the impact of teacher and school characteristics on the relative

probability of a teacher having indexed EVAAS scores that exceed expected growth, the model

was re-written as

$$log \frac{p(Exceeds)}{1-P(Exceeds)} = \beta_1 T + \beta_2 S + fe_d + fe_t + e \qquad (6)$$

where $log \frac{p(Exceeds)}{1-P(Exceeds)}$, represented by variable *IndexExceeds,* is the conditional probability of indexed teacher EVAAS scores exceeding expected growth; T represents a vector of teacher quality and teacher characteristics; S represents a vector of school characteristics; fe_d represents district fixed effects; fe_t represents teacher fixed effects; and e is a standard error term. In this model, T includes variables teacher years of experience, master's degrees, North Carolina Teacher Evaluation Rubric rating, minority status, and gender. S includes variables for school Title 1 status, percentage of non-white students, enrollment, and urbanicity and is integrated into the model to control for the impact of school-level characteristics on indexed EVAAS scores. The fe_d term uses district LEA in order to control for district policies and characteristics within and between school districts that impact teacher quality and student achievement. The fe_t term controls for unobservable teacher characteristics. In testing the model, the researcher also included a fixed effect variable for years to replicate the panel regression environment that controls for teacher variation over time.

Analysis

Multivariate panel regression analyses were run using *xtreg* in Stata 15 to investigate the degree to which years of experience, master's degrees, and North Carolina Teacher Evaluation Rubric rating predicted EVAAS scores. These results informed a response to Research Question 1: In what way, if any, do North Carolina's teacher quality indicators significantly predict students expected growth? Then, controlling for district and year effects, a logistic regression analysis was run using *logistic* in Stata 15 to determine the impact of all teacher variables available, including demographic characteristics, on the relative probability of a teacher having indexed EVAAS scores that exceeded expected growth. These results informed a response to

Research Question 2: What is the profile of the North Carolina teacher whose students exceed expected growth?

In each of these analyses, indexed EVAAS scores served as the dependent variable, while the representative teacher quality indicator(s) or demographic variable(s) served as independent variables. The significance of each regression equation was tested using $a < 10$ and $a < 05$. The researcher used Stata 15 to execute all necessary statistical tests in this study. Spence's Signal Theory, as described in Chapter II, was then used to analyze symmetries and asymmetries found between North Carolina's most popular teacher quality indicators (Research Question 1) and determine the profile of the North Carolina teacher whose students exceed expected growth (Research Question 2) in Chapter V.

Threats to Interpretation

There are number of threats to interpretation that the researcher had to address through modeling procedures. Time was a threat due to changes in teacher effectiveness, school characteristics, and district policies each year. The data was xtset, yielding a panel data set, in order to create a year-by-year fixed effect to control for these time-related threats in Research Question 1. For Research Question 2, where a logistic regression was used, the researcher used a time fixed effect term in the model.

Teacher demographics were a threat because, while those characteristics are unchanged over time, their impact on student performance is unobservable. These teacher effects were controlled for by adding a teacher fixed effect term in each model. Teacher demographic variables were added to the model in Research Question 2 in order to determine their impact on the probability a teacher exceeded expected growth.

School demographics were also a threat because the composition of schools, and school-level policies, can impact student performance. These threats were controlled for by adding school demographic variables in each model. Similarly, the researcher controlled for district effects—stemming from district policies and characteristics within and between school districts that impact teacher quality and student achievement—by adding a district fixed effect term based on district LEA code.

Lastly, the panel data used to investigate Research Question 1 and Research Question 2 was considered unbalanced due to:

- A subset of teachers with incomplete data; and

- A subset of teachers with complete data in individual years that do not appear in every year of the data set.

However, because these instances of missing data are random, the prevailing panel data set is large, the models integrate fixed effects, and Stata applies listwise deletion when running *xtreg* and *logistic* regressions, estimates generated by these models should be consistent with balanced panel sets (Allison, 2002; Baltagi, 2005, StataCorp, 2017).

CHAPTER IV: RESEARCH FINDINGS

The purpose of this quantitative study was to investigate the relationship between signals of teacher quality in North Carolina and student performance, and to determine the profile of North Carolina teachers whose students exceed expected growth. The study used secondary data sets containing teacher quality and school and teacher demographic data furnished by local, state, and national agencies. The chapter starts with a review of the research questions, statistical tests used, and descriptive statistics. A presentation of findings for each research question and summary follows.

The research questions for this correlational study and associated statistical procedures are provided in Table 15.

Table 15

Research Questions and Procedures

Research Question	Statistical Procedure
RQ 1. In what way, if any, do North Carolina's teacher quality indicators significantly predict student growth, as measured by EVAAS?	Panel Multivariate Regression
RQ 2. What is the profile of the North Carolina teacher whose students exceed expected growth?	Logistic Regression

Question 1 was designed to determine whether a statistically significant relationship existed, whether combined or individually, between teacher years of experience, master's degrees, North

Carolina Teacher Evaluation Rubric ratings and indexed EVAAS scores. Question 2 sought to determine the profile of the North Carolina teachers whose indexed EVAAS scores exceeded expected growth. All tests reported as statistically significant were based on $a < 05$.

<div align="center">Results</div>

Research Question 1

A multivariate panel regression was conducted on key teacher quality variables as predictors of indexed teacher EVAAS scores, using the baseline model described in Chapter 3 of

$$Index = \beta_1 X + \beta_2 C + \beta_3 S + fe_t + \varepsilon \qquad (2)$$

where *Index* represents indexed teacher EVAAS scores; X represents teacher years of experience; C represents master's degrees; S represents North Carolina Teacher Evaluation Rubric rating; fe represents fixed effects; and ε represents a standard error term.

The model revealed teacher experience and Rubric ratings to be statistically significant predictors of indexed EVAAS scores, with $p = .000$ for both variables. Teacher experience yielded a coefficient of .15, indicating indexed teacher EVAAS scores increased by .15 for every year increase of teacher experience. The rubric rating coefficient was .10, indicating that indexed teacher EVAAS scores increased by .10 when median observation ratings increased by one unit. Master's degrees did not significantly predict indexed EVAAS scores (p=.701). The calculated R^2 value of .00 was very low and indicates that less than 1% of the variance in indexed EVAAS scores can be explained by the model.

In order to control for school level characteristics, the model was revised and re-run as

$$Index = \beta_1 T + \beta_2 S + fe_t + e \qquad (3)$$

where *Index* represents indexed teacher EVAAS scores; T represents a vector of teacher quality variables; S represents a vector of school characteristics; fe_t represents teacher fixed effects; and

ε represents a standard error term. In this model, T includes variables teacher years of experience, master's degrees, and North Carolina Teacher Evaluation Rubric rating. S includes variables for school Title 1 status, percentage of non-white students, and enrollment. The fe_t term controls for unobservable and demographic teacher characteristics (e.g. race, gender) and ε represents a standard error term.

The model revealed teacher experience and Rubric ratings again to be statistically significant predictors of indexed EVAAS scores, with $p = .000$ for both variables. Teacher experience yielded a coefficient of .13, indicating indexed teacher EVAAS scores increased by .13 for every year increase of teacher experience. The Rubric rating coefficient was again .10, indicating that indexed teacher EVAAS scores increased by .10 when median observation ratings increased by one unit. School characteristics Title 1 status and Non-White Rate both yielded statistically significant coefficients. Master's degrees again failed to significantly predict indexed EVAAS scores, while school enrollment also failed to produce a statistically significant relationship. The calculated R^2 value of .00 was again very low, and indicates that less than 1% of the variance in indexed EVAAS scores can be explained by the model.

Building on the second model, the researcher added a district fixed effect term, fe_d, that leveraged district LEA codes to control for district policies and characteristics within and between school districts that impact teacher quality and student achievement. The revised model, described as

$$Index = \beta_1 T + \beta_2 S + fe_t + fe_d + e \qquad (4)$$

was re-run using this district fixed effect term while all other model components remained the same as in the prior model.

This third model revealed teacher experience ($p = .000$) and Rubric ratings ($p = .038$) again to be statistically significant predictors of indexed EVAAS scores. The coefficient for teacher years of experience was stable in all three models, with the third model coefficient of .15 indicating that indexed teacher EVAAS scores increased by .15 for every year increase of teacher experience. The rubric rating coefficient decreased from .10 to .05 in the third model as district fixed effects were added to the third model, indicating that indexed teacher EVAAS scores increased by .05 when median observation ratings increased by one unit. School non-White rate again yielded a statistically significant relationship, increasing 160% once district fixed effects were added in the third model, and yielded a coefficient of -.90 and indicated that indexed teacher EVAAS scores decreased by .90 for every 1% increase in school non-White rate.

Teacher master's degrees, school Title 1 status, and school enrollment each failed to produce statistically significant relationships. The coefficient for master's degrees was stable in all three models. After yielding significant relationships in the second model, both Title 1 status and school enrollment failed to produce significant relationships in the final model after district effects where added. Further, school Title 1 status saw a 50% coefficient decrease in the final model. The school enrollment coefficient was stable at .00, indicating it had no impact on indexed EVAAS scores.

The calculated adjusted R^2 value of .01 was low but increased compared to the second model and indicates that 1% of the variance in indexed EVAAS scores can be explained by the final model. These results were tested with a clustered error term to further address nesting. Table 16 presents the significant and insignificant coefficients of all three models.

Table 16

Summary of Regression Analyses for Variables Predicting Indexed EVAAS Scores

Variable	Model 1 β	Model 1 SD	Model 2 β	Model 2 SD	Model 3 B	Model 3 SD
Master's Degrees	-.02	.06	-.04	.06	-.02	.08
Evaluation Rubric Rating	0.10**	.02	.10**	.02	.05**	.02
Years of Experience	.15**	.01	.13**	.01	.15**	.01
School Title 1 Status			-.16**	.05	-.08	.06
School Enrollment			.00*	.00	.00	.00
School Non-White Rate			-.34**	.09	-.90**	.20
Adjusted R²	.00		.00		.01	
F	142.12**		78.52**		4.46**	

Note: *p < .10. **p < .05.

Research Question 2

A logistic regression was conducted to determine the impact of all teacher variables available, including demographic characteristics, on the relative probability of a teacher having indexed EVAAS scores that exceeded expected growth. To that end, the final model used in Research Question 1 was revised as

$$log \frac{p(Exceeds)}{1-P(Exceeds)} = \beta_1 T + \beta_2 S + fe_d + fe_t + e \qquad (6)$$

where $log \frac{p(Exceeds)}{1-P(Exceeds)}$, represented by *IndexExceeds,* is indexed teacher EVAAS scores exceeding expected growth (i.e. values greater than or equal to 2.0); T represents a vector of teacher quality and teacher demographic variables; S represents a vector of school

characteristics; fe_d represents district fixed effects; fe_t represents teacher fixed effects; and ε represents a standard error term.

In order to determine which teacher quality, teacher characteristic, and school characteristic variables would be included in T and S, the researcher performed t-tests comparing mean differences between teachers whose indexed EVAAS scores exceeded expected growth (i.e. *IndexExceeds* = 1) and those who did not (i.e. *IndexExceeds* = 0) on available teacher quality and teacher and school characteristics variables. All available variables were revealed to be statistically significant and were therefore included in the model. Table 17 provides a summary of resulting t-tests of teacher quality, teacher characteristic, and school characteristic variables therefore included in the logistic regression.

Table 17

Results of T-tests for Teacher & School Variables by IndexExceeds

	IndexExceeds					
	Does Not Exceed		Exceeds			
	Mean	n	Mean	n	t	df
Teacher Experience	11.12	121,778	11.76	27,502	-11.03**	41,591
Evaluation Rubric Ratings	3.64	98,162	3.87	22,620	-47.68**	34,301
Master's Degrees	.38	125,165	.39	28,229	-5.88**	41,633
Teacher Gender	.81	125,441	.78	28,332	8.99**	40,869
Teacher Minority Status	.19	126,062	.15	28,475	16.01**	45,209
School Title 1 Status	.87	126,119	.79	28,482	31.75**	37,661
School Enrollment	766.45	119,895	883.50	27,186	-34.16**	35,562
School Non-White Rate	.51	119,884	.48	27,183	21.26**	41,884
School Urbanicity	2.52	120,002	2.42	27,221	11.15**	40,994

*Note: *$*p < .10.$ *$**p < .05.$

The subsequent logistic regression model revealed statistically significant relationships for teacher experience ($p = .009$), rubric ratings ($p = .000$), teacher gender ($p = .000$), teacher minority status ($p = .000$), school Title 1 status ($p = .000$), school non-White rate ($p = .000$), and school enrollment ($p = .000$). Teacher master's degrees ($p = .165$) and school urbanicity ($p = .160$) did not yield statistically significant results. These results were tested for interaction effects (e.g. School Title 1 Status and Teacher Minority Status) and none were found.

Teacher experience yielded an odds ratio of 1.00, indicating that single year increases in teacher experience would not improve a teacher's probability of exceeding expected growth. North Carolina Teacher Evaluation Rubric ratings yielded an odds ratio of 1.66, indicating that every unit increase in Rubric rating (e.g. from 1 to 2) would increase a teacher's probability of exceeding expected growth by 66%. Teacher gender yielded an odds ratio of .90, indicating that female teachers were 10% less likely to exceed expected growth than their male counterparts. Teacher minority status yielded an odds ratio of .87, indicating that racial and ethnic minority teachers were 13% less likely to exceed expected growth. School Title 1 status yielded an odds ratio of .83, indicating that teachers at Title 1 schools were 17% less likely to exceed expected growth. School Non-White rate yielded an odds ratio of .76, indicating that every percent increase in Non-White students resulted in a 24% decrease in probability that a teacher would exceed expected growth. School enrollment yielded an odds ratio of 1.00, indicating that increases or decreases in student enrollment would not impact a teacher's probability of exceeding expected growth. The calculated adjusted R^2 value of .05 is low, and indicates that only 5% of the variance in indexed EVAAS scores that exceed expected growth can be explained by the model. Table 18 provides a visual summary of variable probabilities for indexed EVAAS scores exceeding expected growth.

Table 18

Logistic Regression Analysis Summary & Variable Probabilities for Indexed EVAAS Scores Exceeding Expected Growth

	Odds Ratio	SD
Years of Experience	1.00**	.00
Evaluation Rubric Rating	1.68**	.02
Master's Degrees	.98	.02
Teacher Gender	.90**	.02
Teacher Minority Status	.87**	.02
School Title 1 Status	.83**	.02
School Non-White Rate	.76**	.05
School Enrollment	1.00**	.00
School Urbanicity	.99	.01
Adjusted R^2	.05	

Summary

To determine the impact of teacher quality variables on EVAAS scores as described in Research Question 1, the researcher first used a panel multivariate regression that combined teacher quality variables with a fixed effects term to control for unobservable characteristics. In order to develop a more robust model based on these initial findings, the research further derived this initial model to control for the impact of teacher, school, and district effects on teacher quality and student achievement. Of the variables investigated, only two teacher quality variables, years of experience and North Carolina Teacher Evaluation Instrument Rubric rating, and school variable School Non-White Rate produced statistically significant relationships consistently in each model ($p < .05$).

To determine the impact of all teacher variables available, including demographic characteristics, on the relative probability of a teacher having indexed EVAAS scores that exceeded expected growth as described in Research Question 2, the researcher used a logistic regression that combined all available teacher and school variables with teacher, school, district, and year fixed effects terms. Of the variables investigated, teacher experience, North Carolina Teacher Evaluation Rubric ratings, teacher gender, teacher minority status, school Title 1 status, school Non-White rate, and school enrollment produced statistically significant relationships in the model ($p < .05$). Chapter V will present analysis, discussion, and implications for future research and practice based on the findings presented in Chapter IV.

CHAPTER V: CONCLUSIONS AND DISCUSSION

This chapter will present a review of the study and has been divided into four sections. The first section will provide an overview and review the purpose of the study, the second will present conclusions related to the findings from Chapter IV, and the third a discussion of the findings. Lastly, the fourth section will provide implications for future research and recommendations for educational leaders.

Overview and Purpose of the Study

Like many others, the teacher labor market is one where prospective employees and employers send, receive, and evaluate signals that are intended to demonstrate underlying teacher quality. For that reason, Spence's (1973) Signal Theory was used as the conceptual framework for this study, as the model described the ways potential employees tried to differentiate themselves from others in the labor pool to prospective employers. Because teachers all have access to the same set of signals (i.e. higher education, experience), the equilibrium where teachers and prospective employers exchange signals has become disrupted and yields informational asymmetry (Stiglitz, 2002). This is especially true for signals like master's degrees and years of experience, which are pervasive in education labor markets today and can vary considerably in quality despite equivalences in the names of degrees earned and time served, respectively.

Given North Carolina's ongoing need to identify quality teachers in a marketplace wherein candidate characteristics are consistently and finitely signaled, the purpose of this study was to investigate the relationship between those signals of teacher quality in North Carolina—

number of years of experience, master's degrees, North Carolina Teacher Evaluation Rubric ratings—and student performance, measured by EVAAS scores. Specifically, the goal was to determine signal symmetries and asymmetries between the state's outcome measure and the available teacher quality indicators, while providing relevant context on the influence of student and teacher demographics to that end. That is, are these signals of underlying teacher quality accurate predictors of teacher excellence? The specific questions to be answered were:

1. In what way, if any, do North Carolina's teacher quality indicators significantly predict student growth, as measured by EVAAS?

2. What is the profile of the North Carolina teacher whose students exceed expected growth?

The study used secondary data sets containing teacher quality and school and teacher demographic data furnished by local, state, and national agencies for the 2010-2011 through 2013-2014 school years. Multiple regression analyses were run in Stata 15 to investigate relationships between teacher and school variables and informed answers to the aforementioned research questions.

Conclusions

The first research question aimed to determine what relationship, if any, existed between North Carolina's most common teacher quality indicators and student growth, as measured by EVAAS scores. The second research question sought to determine the profile of the North Carolina teacher whose students exceed expected growth by calculating the conditional probability of a teacher exceeding expected growth based on various teacher and school characteristics.

Teacher Characteristics

As described in Chapter IV, teacher variables of years of experience and North Carolina Teacher Evaluation Rubric rating consistently yielded a positive significant relationship in the model for Research Question 1 even as controls for teacher, school, and district effects were added to the model to generate more conservative estimates. Teacher master's degrees consistently failed to yield a significant relationship with EVAAS scores. For Research Question 2, teacher years of experience and North Carolina Teacher Evaluation Rubric ratings again yielded significant relationships; however, only the Rubric rating variable yielded an odds ratio that indicated a positive impact on teacher performance probability. The power of Rubric ratings to determine EVAAS performance in both Research Question 1 and Research Question 2 is important as it suggests that the Rubric—representing a locally derived framework for best pedagogical practice—can differentiate between teachers across the quality spectrum in addition to determining those of highest quality. Teacher gender and minority status also produced significant relationships and yielded odds ratios suggesting a negative impact on performance probability, indicating inequities in North Carolina's educational system and institutions had an impact on teacher performance metrics. Teacher master's degrees again failed to yield a significant relationship in the model for Research Question 2.

These findings are consistent with previous research completed by those that have argued for teacher experience as an effective predictor of student performance (Dewey, Husted, & Kenny, 2000; Greenwald, Hedges, & Laine, 1996; Nye, Konstantopoulos, & Hedges, 2004), particularly those using longitudinal (panel) data (Clotfelter, Ladd, & Vigdor, 2007; Croninger, Rice, Rathbun, & Nisho, 2007; Kane, Rockoff, & Staiger, 2008). These findings also seem to confirm the body of research suggesting the lack of impact that teacher master's degrees have on

student performance that has been championed by Hanushek (1986, 1989, 1991, 1997) and others (Aaronson, Barrow, & Sanders, 2007; Chingos & Peterson, 2011; Goldhaber & Brewer, 1997; Ladd, & Vigdor, 2007; Kane & Staiger, 2012; Koedell & Betts, 2007; Rice, Rathbun, & Nisho, 2007). Lastly, these findings seem to be consistent with previous research that found significant relationships between teacher observations and student achievement (Goldhaber & Brewer, 1997; Murnane & Phillips, 1981; Tyler, Taylor, Kane, & Wooten, 2010), including value-added scores (Bill & Melinda Gates Foundation, 2010; Cantrell & Kane, 2013; Dee & Keys, 2004; Jacob & Lefgren, 2005; Kane & Staiger, 2012; Sanders & Horn, 1994). Beyond confirming findings in the research base, these finding add credence to the inclusion of teacher experience and the evaluation rubric in North Carolina's teacher evaluation system while contributing further evidence to the removal of master's degrees from compensation plans.

The researcher can also make conclusions about the profile of the North Carolina teacher whose students exceeded expected growth by interpreting odds ratios associated with each variable. To that end, the researcher can reasonably conclude that teachers who were White, male, had higher North Carolina Teacher Evaluation Rubric ratings, and taught at non-Title 1 schools with smaller minority student populations had the highest collective probability of exceeding expected growth. These findings are significant given the teacher workforce in North Carolina is overwhelmingly female, its student population increasingly non-White, and the prevalence of Title 1 schools across the state. Taken together with the impact that teacher race—namely, Whiteness—had on student performance, these finding may provide evidence of flaws in state, district, and school policies and systems that disproportionately impact women, students from low socioeconomic backgrounds, and persons of color. Further, these findings make current efforts to recruit, evaluate, and retain teachers less effective and more complex.

School and District Characteristics

Higher school non-White populations consistently yielded a negative, significant relationship in the model for Research Question 1 and saw a coefficient increase of over 160% once district level effects were controlled for in the model. This is an important finding as it suggests decreasing racial minority student populations is related to increases in student growth. It is unclear whether this finding is a byproduct of recent school districting policies in North Carolina that have created increasingly segregated neighborhood schools, or school policies such as student tracking that make it more difficult for students of color to excel on state assessments by disproportionately funneling them into classrooms with less rigorous curriculum, instruction, and experiences (Andre-Bechely, 2005; Ansalone, 2006; Kahne, 1994; Kelly, 2007; Marshall & Gerstl-Pepin, 2005). School-level variables related to Title 1 status and school enrollment initially yielded significant relationships but failed to produce significance once district level effects were controlled for in the model. These findings suggest that socioeconomic status—a hallmark of recent school zoning policies—may not have an impact on student performance in North Carolina's EVAAS growth model. Similarly, these findings also suggest that school type had no effect on student performance and indicates stability across the K-12 spectrum despite the variety and number of tests given at the high school level.

For Research Question 2, Title 1 status, non-White rate, and enrollment yielded significant relationships. However, only odds ratios for Title 1 status and school non-White rate indicated an impact, albeit negative, on performance probability and suggests that teachers at affluent schools with lower non-White student populations had a higher probability of exceeding expected growth. School urbanicity failed to yield a significant relationship, indicating that resource disparities and cultural differences between urban, suburban, and rural communities

may not impact student growth in the North Carolina EVAAS model. Taken in total, these findings provide further evidence that some school effects are associated with teacher effectiveness and suggests that school leadership and culture have a significant impact on the ability of students to exceed expected growth.

Variation Explained by Modeling

Given the model(s) used by the researcher at best produced adjusted R^2=.01 for Research Question 1, it is reasonable to conclude that the preponderance of the variation in teacher quality—and EVAAS scores, in particular—cannot be explained by common teacher quality variables used in North Carolina. Similarly, the model used by the researcher to answer Research Question 2 produced adjusted R^2=.05, so it is also reasonable to conclude that the preponderance of the variation related to the probability that a teacher exceeds expected growth in North Carolina cannot be explained by the variables available or commonly used in teacher quality research. These findings suggest that diversifying measurement types does not necessarily yield a complete picture of teacher effectiveness, which is important given the proclivity of educational leaders and policymakers to make data driven decisions that are then scaled across varied educational contexts.

<div align="center">

Discussion and Implications

</div>

While teacher quality research on the efficacy of teacher years of experience, master's degrees, and teacher observations in predicting student achievement has been completed for decades, this study represents a contribution to the research base by incorporating North Carolina specific parameters—namely, the North Carolina Teacher Evaluation Rubric and EVAAS—and a conceptual lens in Spence's (1973) Signal Theory that focused on validating the accuracy of signals used to communicate the underlying quality of teacher excellence. These are important

distinctions, as the researcher aimed to inform educational leaders interested in identifying

excellent teaching in North Carolina in order to improve the effectiveness of their hiring,

compensation, retention, and policymaking efforts while providing additional context on the

influence of teacher and student demographics on teacher performance indicators. The sections

that follow discuss the implications of these findings, specifically as they relate to the existing

research base, the underlying assumptions of North Carolina's teacher evaluation system, and the

model used in this study.

Connecting North Carolina to the Nation

Concluding that teacher years of experience and higher quality pedagogy (i.e. teachers

who score higher on the evaluation rubric) positively impact student achievement are expected

outcomes given the research base and common policy logic. Similarly, the failure of master's

degrees—and advanced degrees more broadly—to predict student performance is also a

conclusion supported by the research base and increasingly popular in policy discourse as

evidenced by North Carolina's decision to no longer compensate teachers for them (Banchero &

Rutland, 2013). To that end, increased compensation tied to master's degrees could be

considered more as a cost of living adjustment—or even a financial incentive for prospective

teachers without undergraduate degrees in education to pursue teaching careers—than a teacher

quality variable. Whatever the corrective policy action, the prevalence of master's degrees in the

teacher labor market, with 37% of those studied having an advanced degree, will make

modifying its value as a signal of teacher quality a difficult practical and political proposition for

education policymakers.

Unlike teacher experience and master's degrees, there is no prior published teacher

quality research that included teacher observations (i.e. Rubric ratings) and value-added scores

(i.e. EVAAS) unique to North Carolina. To that end, these findings add credibility to the locally

created North Carolina Teacher Evaluation Rubric as that variable consistently outperformed

others in predicting EVAAS scores and yielded a 66% increase in probability of exceeding

expected growth for every one unit increase on the Rubric. Further testing by the researcher

revealed that 13% of the variation in rubric ratings could be explained by teacher years of

experience and master's degree attainment, both of which yielded positive significant

relationships. Table 19 provides a visual summary of the variables predicting rubric ratings.

Table 19

Summary of Regression Analyses for Variables Predicting Rubric Ratings

Variable	β	SD
Educational Attainment	.06**	.02
Years of Experience	.07**	.00
Title 1 Status	-.00	.01
School Enrollment	-.00	
School Non-White Rate	-.06	.05
Adjusted R^2	.13	

Note: *p < .10. **p < .05.

Of note, school Title 1 status, enrollment, and non-White rate had no impact on Rubric

ratings, which would indicate that teacher observations are a more appropriate mechanism for

teacher evaluation than critics decrying its subjectivity, sensitivity to student assignment, and

lack of fidelity would suggest (Amrein-Beardsley, 2008; Goldrick, 2002; Hanushek, 1986; Hazi

& Rucinski, 2009; Kupermintz, 2003; Lefkowitz, 2000; Riley, 1985-1986; Varma & Stroh,

2001; Wayne & Ferris, 1990). These findings also suggest that the Rubric seems to control well

for the inequities in North Carolina's educational system and institutions in ways that other

teacher quality indicators—and the evaluation and compensation systems more broadly—do not.

Thus, while not directly related to compensation in North Carolina, these findings could open up

opportunities for the Rubric to support more reliable recruitment, hiring, and retention policies

and practices at the school and district level in the state.

A System with Unfulfilled Promises

While North Carolina connections and context related to existing research are helpful to

educational leaders, the researcher also finds critical flaws in the underlying assumptions

inherent in the teacher evaluation system. As presented in Chapter 2, the central benefit to value

added models—including North Carolina's EVAAS model—is an alleged leveling of the playing

field for students and teachers that are marginalized under a model rewarding proficiency.

Student growth, where individualized targets for improvement replace fixed thresholds that

indicate a passing score, have been described to educators and the general public by education

reformers as a fairer way to determine student performance and teacher quality. Indeed, the

antecedents of equity are evident in North Carolina's EVAAS model when it purportedly

accounts for immutable student characteristics (e.g. race) and even uses prior student

performance on a variety of state assessments (e.g. using English Language Arts scores in

predicting expected Algebra 1growth) to create as accurate a growth prediction as possible. What

resulted from the implementation of EVAAS was a promise to all students and teachers in North

Carolina that their efforts would be equitably measured and rewarded. This promise has been

unfulfilled.

For teachers, the evidence of this broken promise is most directly seen in the findings

associated with Research Question 2, where the researcher concluded that teachers who were

White, male, and had higher median North Carolina Teacher Evaluation Rubric ratings had the highest probability of exceeding expected growth. This is a remarkable finding given that, outside of Rubric ratings, North Carolina teachers seem to have little control of their ability to demonstrate teacher excellence as it relates to EVAAS scores. Instead, these findings provide evidence of further marginalization of teachers and students traditionally harmed in proficiency-based systems. This conclusion therefore raises serious concerns about the very foundation of teacher quality, assessment, and evaluation in North Carolina. Otherwise, how can it be explained that female teachers are 10% less likely to exceed expected growth, despite representing 81% of all teachers in North Carolina? Similarly, what does it say about an evaluation system when 18% of the workforce, comprising persons of color teaching the state's children, is less likely to have students who exceed expected growth based on the fact that they are non-White? The answers to these questions are unable to be answered by this research study but are worthy of further investigation as gender and Whiteness should not predispose a teacher or her students to success or failure.

Practically, these findings present major challenges for education leaders and policymakers. Taken at face value, the findings would advise incentivizing the recruitment and retention of teachers hailing from majority populations. This would lead to the further marginalization of students of color in increasingly racially isolated settings as research has shown the negative impact for racial minority students without teachers who look like them (Dee, 2005; Oates, 2003; Sharma, Joyner, & Osment, 2014). More responsibly, educational leaders should take these findings as an opportunity to examine the systems and processes that seem to disproportionately affect the opportunity of female and non-White teachers to exceed expected growth. Further, this examination should include the racialized and gendered

dimensions of teaching in different geographical contexts across the state and the fidelity of culturally responsive policies and instruction.

Outside of teacher demographics, the system as currently constructed has also broken a promise to students in that performance in the EVAAS growth model seems to follow proficiency trends that overwhelmingly reward White and affluent students. In answering Research Question 1, it was found that both Title 1 status and non-White rate yielded significant, negative relationships findings when school variables were added to the regression model. When district fixed effects were added to the model the significance of Title 1 faded, but the impact of schools' non-White rate prevailed. These findings suggest that variation exists across districts with respect to EVAAS scores and points to school effects as a significant contributor to winners and losers in the North Carolina's value-added model. Specifically, these school effects related to student poverty and demographic characteristics, as the researcher concluded in answering Research Question 2 that teachers at non-Title 1 schools (17% increase) with smaller minority student populations (24% increase) had a higher probability of exceeding expected growth.

Like teacher race and gender, student demographics should not impact the relative probability of success in a system built on growth; that is, students from traditionally marginalized groups should not have a more difficult time than their privileged peers when the accountability playing field has been leveled as a result of primarily competing against themselves. What is unclear about these findings are the causes of such a discriminatory outcome, as implications include the possibility that lower quality teachers tend to teach at impoverished schools with high percentages of non-White students, that school policies and procedures are responsible for depressing teacher effectiveness, or some combination of both. It is also worthy to investigate whether policies or cultural components within affluent and

predominately White schools attract objectively better teachers or otherwise promote higher quality instruction. Regardless, in an effort to create a more equitable system of student performance measurement, educational leaders should further investigate and aim to correct for the impact uncontrollable factors such as student poverty and race have on student performance.

In summary, the signals sent, received, and evaluated in the teacher labor market can only be interpreted accurately if the environment in which they are signaled is based on underlying quality and free of bias. The influence of demographics and school effects are representative of unfulfilled promises to teachers and students in pursuit of achieving and exceeding expected growth. To that end, signal asymmetries arising from these findings—caused by employees' and employers' increased difficulty in demonstrating and determining teacher quality, respectfully—will certainly bolster arguments against standardized state assessments and the EVAAS value-added model based on those results.

Macro Analysis and Model Vulnerability

In order to answer Research Question 1 and Research Question 2, the researcher constructed a panel data set that represented teacher quality data from 2010-2011 to 2013-2014. This was done not only to increase the sheer number of observations and improve confidence in subsequent findings, but also to mirror at the macro level the practice in North Carolina of using three-year rolling averages to determine individual teacher EVAAS scores and ratings yielding does not meet, meets, or exceeds expected growth designations. The decision to take a state-level, macro focus was mostly based on the researcher's intent on providing findings, conclusions, and recommendations that would be applicable to educational leaders across the state. This strategy was also employed because limiting the data set to only teachers with three year rolling averages would reduce the number of observations available for the model and

create issues in the modeling process without clear policy-driven solutions, such as instances where three years of non-sequential data existed in the four years studied. Ultimately, the researcher contends that this macro approach provided the greatest contribution to the research base as it provides a starting point for educational leaders interested in replicating similar studies in local school districts where common data issues (e.g. missing data) and teacher employment is presumably more stable.

Despite leveraging a state-level, macro approach that incorporated nearly every available teacher and school variable in the research base, only 1% of the variation in teacher EVAAS scores could be explained by the model used in answering Research Question 1 and only 5% of the variation in teacher exceeds expected growth probability could be explained by the model used in answering Research Question 2. These findings indicate that, despite advances in modeling methodologies and data collection, a significant amount of variation in teacher quality is left to be explained. Educational leaders are in the business of finding policy solutions that treat people equally and are scalable across contexts, so it is fair to speculate that this desire may have driven North Carolina's teacher evaluation process into a system that is wholly imperfect and altogether incomplete. The system is imperfect because winners and losers seem to be predisposed based on uncontrollable characteristics as described in the previous section, and incomplete because between 95% and 99% of teacher quality the researcher simply cannot account for. This should give us all great pause as we grapple with ongoing accountability movements in North Carolina and remain focused on creating a system based on equitable access to success and excellence. Lastly, these findings suggest that traditionally marginalized teachers and students do not seem to fare any better in a system based on growth, so perhaps education reform in North Carolina should move away from a paradigm so inextricably tied to high-stakes

testing. Failing to do so will only further a persistent achievement gap and overall stagnation on national and international assessments. Worse still, the system as currently constructed stands to further erode confidence in public schools, teachers, and education. Maybe more of our reform efforts, after all, should be focused on where it seems to matter most: improving the pedagogical practices of our teachers for the collective benefit of all students.

Recommendations

This study offers findings that suggests a significant, positive relationship exists between teacher quality variables years of experience and North Carolina Teacher Evaluation Rubric ratings and student growth, as measured by EVAAS scores. Upon further investigation of these findings as well as the profile of the teacher whose students exceed expected growth, the researcher concluded that many factors outside of teachers' control—namely, their gender, minority status, school poverty level, and school demographics—seemed to confound teachers' ability to meet expected growth. Further, the models used by the researcher, which integrated nearly all of the teacher, school, and district data available, were able to explain no more than 5% of the variance in EVAAS Scores, which yields wide-ranging implications on North Carolina's approach to teacher quality. To that end, the following sections offer implications for future research and recommendations for educational leaders.

Recommendations for Future Research

Additional research is needed in order to more fully explain and contextualize the findings presented in this study. This study aimed to provide a bridge between the existing research base and North Carolina specific applications. To that end, future research should include, but is not limited to:

1. Investigations that test the findings presented here when teacher quality variables are more fragmented (e.g. early year vs. career teachers, master's degrees vs. advanced degrees);

2. Investigations that explore the impact of subject area (e.g. Math) on the relationship between teacher quality variables and both EVAAS scores and exceeding expected growth probabilities;

3. Studies that more directly examine the impact of school type (i.e. elementary, middle, and high school) on EVAAS scores and exceeding expected growth probabilities; and

4. More detailed analysis on the impact of minority teacher and student populations on EVAAS scores and exceeding expected growth probabilities that accounts for variety in subgroup performance.

Lastly, because of the small amount of variation in EVAAS scores explained by the models used in this study, it may be valuable to begin incorporating emerging data that has not traditionally been used in teacher quality research, like student and parent surveys.

Recommendations for Educational Leaders

This study presented findings that suggested teacher years of experience and the North Carolina Teacher Evaluation Rubric are accurate signals of teacher quality in North Carolina. In practice, educational leaders could use these indicators to further develop recruiting, hiring, compensation, and retention strategies pursuant to improving and sustaining student achievement in their schools and districts. However, educational leaders and policymakers alike should strongly consider the impact of teacher and school demographic characteristics before extrapolating these findings for use in wide ranging policy proposals, as current compensatory and retention policies that do not account for the concerns presented in this chapter could lead to

the further marginalization of vulnerable student populations and discrimination against teachers from protected classes. To that end, it is strongly recommended that school district leaders replicate this study with school and district specific variables of interest in an effort to increase the explanatory power of the model used before making policy decisions and implementations.

APPENDIX A: FULL VARIABLE LIST

Variable	Characteristic	Measurement	Representation
Years of Experience	Teacher	Years	Completed years of service
Master's Degrees	Teacher	Advanced Degree	0 = No 1= Yes
Teacher Evaluation Rubric	Teacher	Median Rating	0 = Not Demonstrated 1 = Developing 2 = Proficient 3 = Accomplished 4 = Distinguished
Index	Teacher	Composite EVAAS Index	Indexed EVAAS Score
Index Exceeds	Teacher	Exceeds Expected Growth	0 = Index Does Not Exceed 1 = Index Exceeds
Gender	Teacher	Male or Female	0 = Male 1 = Female
Minority	Teacher	Non-white	0 = No 1 = Yes
Title 1	School	Status	0 = Title 1 Ineligible 1 = Title 1 Eligible
Enrollment	School	Number of Students	Total Number of Students
Urbanicity	School	Federal Rating	11 – Large City 12 – Medium City 13 – Small City 21 – Suburb Large 22 – Suburb Midsize 23 – Suburb Small 31 – Town Fringe 32 – Town Distant 33 – Town Remote 41 – Rural Fringe 42 – Rural Distant 43 – Rural Remote
Non-White Rate	School	Percentage	Non-White Percentage

PUBLIC SCHOOLS OF NORTH CAROLINA
DEPARTMENT OF PUBLIC INSTRUCTION | June St. Clair Atkinson, Ed.D., *State Superintendent*
WWW.NCPUBLICSCHOOLS.ORG

Memorandum of Agreement
By and Between the North Carolina Department of Public Instruction and the Education Policy Initiative at Carolina

The Education Policy Initiative at Carolina (EPIC) is an interdisciplinary team that conducts rigorous research and evaluation to inform education policy and practice. We produce evidence to guide data-driven decision-making using qualitative and quantitative methodologies tailored to support all education stakeholders. Our work is ultimately driven by a vision of high quality and equitable education experiences for all students in North Carolina.

Data Required

North Carolina policymakers continue to grapple with identifying the essential elements of excellent teaching so that recruitment and compensation policies can appropriately value teacher quality indicators (e.g. teacher credentials) and measures of teacher performance (e.g. EVAAS estimates). Towards this end, there is a lack of research evidence describing the relationships between teacher characteristics and teacher performance. This complicates state and school district efforts to make informed policy decisions and improve student achievement.

Therefore, this study intends to inform educational leaders and policymakers of the indicators that most accurately predict high student growth in North Carolina. The proximal outcome measures for this study include a) an analysis of teacher-level quality data (e.g. Years of Experience, Master's Degrees, NCEES ratings) and demographic data (e.g. gender, race/ethnicity, school poverty designation) and b) teacher-level EVAAS ratings for the purpose of answering the following evaluation questions:

1. Do any North Carolina teacher quality indicators, or combination of indicators, meaningfully predict whether a teacher's students exceeded expected growth?
2. What is the profile of the North Carolina teacher whose students exceed expected growth?

These analyses are to be completed by Jason Sinquefield, a doctoral candidate at UNC Chapel Hill and an affiliate at the EPIC. For these analyses we request the authority to use NCDPI data already housed at the EPIC. These data include teacher EVAAS estimates, teacher evaluation ratings, teacher demographics, teacher years of experience, and teacher educational attainment for the 2009-10 through 2013-14 academic years. Upon completion of these analyses, a white paper will be provided to NCDPI pursuant to TERMS 6 and 7.

Agreement for Sharing of Data

This Agreement is entered into by the NCDPI and the EPIC for the purpose of sharing information between the parties in a manner consistent with the Family Education Records Privacy Act of 1974 ("FERPA"). The information will be used by researchers at the EPIC to conduct studies designed to improve instruction for children in the state of North Carolina.

The Family Educational Rights and Privacy Acts Statute (FERPA) describes circumstances under which Local Educational Agencies (LEAs) and the NCDPI are authorized to release confidential data regarding individual students, teachers, and schools without prior parental consent. Confidential information can be disclosed to organizations as stated in section II. 2. and is destroyed per section V.

Consistent with NCDPI nomenclature, EPIC agrees to ensure that all reports (using data from the National School Lunch Program) will use the terms "Economically Disadvantaged Students" as opposed to "Students that qualify for Free/Reduced Price meals" if the data are reported by sub-groups.

The following terms further specify the manner in which the NCDPI agrees to share data with the EPIC, subject to FERPA regulations:

I. PARTIES. The NCDPI is a state educational authority authorized to receive information from local educational agencies ("LEAs") subject to FERPA, as authorized by 34 CFR Section 99.31. Researcher desires to conduct studies on behalf of the NCDPI and LEAs for the purpose of improving instruction in North Carolina public schools. The parties wish to share data collected by the NCDPI regarding education in North Carolina, some of which may allow the identification of individual students.

II. COMPLIANCE WITH FERPA. To effect the transfer of data subject to FERPA, the EPIC agrees to:

1. In all respects comply with the provisions of FERPA. For purposes of this Agreement, "FERPA" includes any amendments or other relevant provisions of federal law, as well as all requirements of Chapter 99 of Title 34 of the Code of Federal Regulations and reauthorization when effective. Nothing in this Agreement may be construed to allow either party to maintain, use, disclose or share student information in a manner not allowed by federal law or regulation.

2. Use the data shared under this Agreement for no purpose other than research and analysis authorized under Section 99.31(a)(6) of Title 34 of the Code of Federal Regulations which allow disclosure of personally identifiable information from students' education records in connection with NCDPI's conducting studies to develop, validate, or administer predictive tests, administer student aid programs, or improve instruction. Use the data pursuant to 42 USC 1758(b)(6)(A)(i)(II) which allows for disclosure of student eligibility for free or reduced priced meals under the USDA's School Lunch Program for a State education program administered by the State or local educational agency. The EPIC further agrees not to share data received under this MOA with any other entity

without the NCDPI approval. The EPIC agrees to allow the Office of the State Auditor, subject to FERPA restrictions, access to data shared under this Agreement and any relevant records of the EPIC for purposes of completing authorized audits of the parties.

3. Require all employees, contractors and agents of any kind to comply with all applicable provisions of FERPA and other federal laws with respect to the data shared under this Agreement. The EPIC agrees to require and maintain an appropriate confidentiality agreement from each employee, contractor or agent with access to data pursuant to this Agreement. Nothing in this paragraph authorizes sharing data provided under this Agreement with any other entity for any purpose other than completing the EPIC's work authorized under this Agreement.

4. Maintain all data obtained pursuant to this Agreement in a secure computer environment and not copy, reproduce or transmit data obtained pursuant to this Agreement except as necessary to fulfill the purpose of the original request. All copies of data of any type, including any modifications or additions to data from any source that contains information regarding individual students, are subject to the provisions of this Agreement in the same manner as the original data. The ability to access or maintain data under this Agreement shall not under any circumstances transfer from the EPIC to any other institution or entity or unauthorized individual or agent.

5. Not to disclose any data obtained under this Agreement in a manner that could identify an individual student, except as authorized by FERPA, to any other entity. The EPIC may publish results of studies authorized by this Agreement, but specifically agrees to abide by the NCDPI "small numbers" policy of deleting all data items that include any group of students less than five (5), and to require all employees, contractors and agents of any kind to also abide by that policy.

6. Not to provide any data obtained under this Agreement to any party ineligible to receive data protected by FERPA or prohibited from receiving data from any entity by virtue of a finding under Section 99.31(6)(iii) of Title 34, Code of Federal Regulations.

7. Provide to the NCDPI a list of specific research studies, updated semi-annually, for which the confidential data are being used, and to notify the NCDPI in advance and in writing of any new project or research question the EPIC proposes to address. This list of research studies will identify linkages of all data possessed by the EPIC under this Agreement and covered by FERPA to specific research studies. Further, it will include the fixed ending date for use of all data linked to each project. EPIC agrees to neither amend nor alter the scope, design, format or description of a project or report generated by the EPIC for this project, except as consistent with the Agreement, without prior written notice to NCDPI.

8. Destroy all data and provided verification in writing of the destruction of all copies of the data obtained under this Agreement to NCDPI 12 months following the date of publication of the final report of this project. All data no longer needed shall be destroyed or returned to the NCDPI in compliance with 34 CFR Section 99.35(b)(2). The EPIC

agrees to require all employees, contractors, or agents of any kind to comply with this provision.

III. DATA REQUESTS. The NCDPI may decline to comply with a request if it determines that providing the data requested would not be in the best interest of North Carolina public schools. All requests shall include a written statement of the purpose for which it is requested and an estimation of the time needed to complete the project for which the data is requested. Data requests may be submitted in writing by post, electronic mail or facsimile.

IV. AUTHORIZED REPRESENTATIVE. The EPIC shall designate in writing a single authorized representative able to request data under this Agreement. The authorized representative shall be responsible for transmitting all data requests and maintaining a log or other record of all data requested and received pursuant to this Agreement, including confirmation of the completion of any projects and the return or destruction data as required by this Agreement. The EPIC designates *Kevin Bastian* as the authorized representative of the Department's data. NCDPI or its agents may upon request review the records required to be kept under this section.

V. RELATED PARTIES. The EPIC represents that it is authorized to bind to the terms of this agreement, including confidentiality and destruction or return of student data, all related or associated institutions, individuals, employees or contractors who may have access to the data or may own, lease or control equipment or facilities of any kind where the data is stored, maintained or used in any way. Data may be stored on a server with additional data but may not be merged with any other data without prior written permission from NCDPI. This Agreement takes effect only upon acceptance by authorized representatives of the EPIC, by which that institution agrees to abide by its terms and return or destroy all student data covered by this MOA 12 months following the date of publication of the final report of this project.

TERMS.

1. This Agreement takes effect upon signature by the authorized representative of each party and will remain in effect until December 31, 2018. The parties further understand that the NCDPI may cancel this Agreement at any time, upon reasonable notice. The NCDPI specifically reserves the right to cancel this Agreement should the NCDPI, in its sole discretion, determine that confidential student information has been released in a manner inconsistent with this Agreement, has not been maintained in a secure manner, or that substantially similar data access has become generally available for research purposes through any other mechanism approved by the NCDPI.

2. The EPIC understands that the Agreement does not convey ownership of data to the EPIC.

3. Any cloud storage or processing will require the express written consent of NCDPI.

4. Data from NCDPI must not be taken outside the United States.

5. The EPIC will provide NCDPI with an electronic copy of the program code created by the EPIC to

perform analysis within its statistical software platform (e.g., .do files created for use in Stata). NCDPI agrees to keep such files and a copy of the original data for at least 7 years.

6. EPIC will provide NCDPI with an electronic copy of the final versions of all reports and other documents associated with the project. NCDPI, as the owner of the data, reserves the right to distribute and otherwise use the final report and associated documents in its discretion, in sum or in part. The EPIC, or its agents working on this project, retain the right to publish findings in other publications, provided that prior notice of report is first shared with NCDPI.

7. The EPIC has the right, consistent with scientific standards, to publish, present or use the study results gained in the course of the research under this Agreement. In order to protect the confidentiality of previously identified confidential information disclosed to the EPIC, the authorized representative agrees to provide to NCDPI any proposed publications or presentations which are to make public any findings, data, or results of the research under this Agreement for the Department's review at least thirty (30) days prior to submission of a manuscript or abstract for publication or the date of the presentation. The EPIC agrees to delete any of NCDPI's previously identified confidential information therefrom.

Entered into this _____ day of _____, 2016.

Accepted on behalf of the North Carolina Department of Public Instruction

By _____
June St. Clair Atkinson, State Superintendent
North Carolina Department of Public Instruction

Accepted on behalf of EPIC

By _____

By _____
Authorized Representative

97

CONFIDENTIALITY AGREEMENT

I am working as an agent for the state of North Carolina on a project involving North Carolina schools, students, teachers, and/or administrators. I understand that my work on this project involves the use of Department of Public Instruction or Local Education Agency (LEA) data that is confidential under state law, federal law, or both state and federal law.

All personally identifiable information is to be protected in adherence with FERPA guidelines. I will refrain from including personally identifiable information in any form of communication with anyone outside the project. This includes emails, instant messaging, faxes, other written correspondence, and any type of oral conversation. When conversing with any LEA or school employees about any students, staff, schools, or LEAs in the execution of my assigned duties, I will take all precautions to protect the confidentiality of all personally identifiable information.

I understand that I can be removed from this project if it is determined that I either intentionally violated or was willfully negligent on any aspect of the Confidentiality Agreement. Further, my violation of or negligence regarding this Confidentiality Agreement may put in jeopardy the working relationship between my agency and the Department of Public Instruction. I also understand that my violation of this Confidentiality Agreement could result in my being held liable for damages in a civil lawsuit.

Name (please print): _____

Signature: _____

Date: _____

COMPLIANCE WITH FERPA. To effect the transfer of data subject to FERPA, Contactor agrees to:

1. In all respects comply with the provisions of FERPA. For purposes of this agreement, "FERPA" includes any amendments or other relevant provisions of federal law, as well as all requirements of Chapter 99 of Title 34 of the Code of Federal Regulations. Nothing in this agreement may be construed to allow either party to maintain, use, disclose or share student information in a manner not allowed by federal law or regulation.

2. Use the data shared under this agreement for no purpose other than work authorized under Section 99.31(a)(6) of Title 34 of the Code of Federal Regulations. Agent further agrees not to share data received under this MOU with any other entity without the NCDPI approval. Agent agrees to allow the Office of the State Auditor, subject to FERPA restrictions, access to data shared under this agreement and any relevant records of Agent for purposes of completing authorized audits of the parties.

3. Maintain all data obtained pursuant to this agreement in a secure computer environment and not copy, reproduce or transmit data obtained pursuant to this agreement except as necessary to fulfill the purpose of the original request. All copies of data of any type, including any modifications or additions to data from any source that contains information regarding individual students, are subject to the provisions of this agreement in the same manner as the original data. The ability to access or maintain data under this agreement shall not under any circumstances transfer from Agent to any other institution or entity.

4. Not to disclose any data obtained under this agreement in a manner that could identify an individual student, except as authorized by FERPA, to any other entity. Agent agrees to abide by the NCDPI "small numbers" policy of deleting all data items that include any group of students less than five (5).

5. Not to provide any data obtained under this agreement to any party ineligible to receive data protected by FERPA or prohibited from receiving data from any entity by virtue of a finding under Section 99.31(6)(iii) of Title 34, Code of Federal Regulations.

6. Destroy all data obtained under this agreement when it is no longer needed for the purpose for which it was obtained. Nothing in this agreement authorizes the Agent to maintain data beyond the time period reasonably needed to complete the purpose of the request. All data no longer needed shall be destroyed or returned to the NCDPI in compliance with 34 CFR Section 99.35(b)(2).

PERSONALLY IDENTIFIABLE INFORMATION (PII): Any information about an individual maintained by an agency, including any information that can be used to distinguish or trace an individual's identity such as name, social security number, date or place of birth, mother's maiden name, biometric records, and any other personal information that is linked or linkable to an individual.

APPENDIX C: SUMMARY OF INDEX EVAAS COMPOSITE SCORE CALCULATION

The process for creating teacher-level composite index scores using EVAAS data is described

below:

1. Calculate a composite EVAAS gain score using the formula below:

$$Composite\ Gain = \frac{1}{n}(Score_1) + \cdots + \frac{1}{n}(Score_n)$$

   Where *n* is the number of EVAAS scores available and *Score* represents the reported

   EVAAS gain for each assessment for a given teacher in a given year.

2. Calculate a composite EVAAS standard error using the formula below:

$$Composite\ SE = \frac{1}{n}\sqrt{(Score_1 SE)^2 + \cdots (Score_n SE)^2}$$

   Where *n* is the number of EVAAS scores available and *Score* represents the reported

   EVAAS standard error for each assessment for a given teacher in a given year.

3. Calculate a composite EVAAS Index using the formula below:

$$Composite\ Index = \frac{Composite\ Gain}{Composite\ SE}$$

4. Standardize Composite Index scores so that each score has a standard error of 1.

APPENDIX D: NORTH CAROLINA TEACHER EVALUATION RUBRIC

Rubric for Evaluating North Carolina Teachers
(Required for Self-Assessment and for Observations)

This form should be used for the teacher self-assessment, classroom observation, and the summary evaluation.

Name: _____   Date: _____

School: _____   District: _____

Evaluator: _____   Title: _____

Start Time: _____   End Time: _____

Standard I: Teachers Demonstrate Leadership

Element Ia. Teachers lead in their classrooms. Teachers demonstrate leadership by taking responsibility for the progress of all students to ensure that they graduate from high school, are globally competitive for work and postsecondary education, and are prepared for life in the 21st century. Teachers communicate this vision to their students. Using a variety of data sources, they organize, plan, and set goals that meet the needs of the individual student and the class. Teachers use various types of assessment data during the school year to evaluate student progress and to make adjustments to the teaching and learning process. They establish a safe, orderly environment, and create a culture that empowers students to collaborate and become lifelong learners.

Developing	Proficient	Accomplished	Distinguished	Not Demonstrated (Comment Required)
	… and	… and	… and	
❑ Understands how they contribute to students graduating from high school.	❑ Takes responsibility for the progress of students to ensure that they graduate from high school.	❑ Communicates to students the vision of being prepared for life in the 21st century.	❑ Encourages students to take responsibility for their own learning.	
❑ Uses data to understand the skills and abilities of students.	❑ Provides evidence of data-driven instruction throughout all classroom activities.	❑ Evaluates student progress using a variety of assessment data.	❑ Uses classroom assessment data to inform program planning.	
	❑ Establishes a safe and orderly classroom.	❑ Creates a classroom culture that empowers students to collaborate.	❑ Empowers and encourages students to create and maintain a safe and supportive school and community environment.	

Element Ib. Teachers demonstrate leadership in the school. Teachers work collaboratively with school personnel to create a professional learning community. They analyze and use local, state, and national data to develop goals and strategies in the school improvement plan that enhances student learning and teacher working conditions. Teachers provide input in determining the school budget and in the selection of professional development that meets the needs of students and their own professional growth. They participate in the hiring process and collaborate with their colleagues to mentor and support teachers to improve the effectiveness of their departments or grade levels.

	. . . and	. . . and	. . . and	
❑ Attends professional learning community meetings.	❑ Participates in professional learning community.	❑ Assumes a leadership role in professional learning community.	❑ Collaborates with colleagues to improve the quality of learning in the school.	
❑ Displays awareness of the goals of the school improvement plan.	❑ Participates in developing and/or implementing the school improvement plan.	❑ Collaborates with school personnel on school improvement activities.	❑ Assumes a leadership role in implementing school improvement plan throughout the building.	

101

	Developing	Proficient	Accomplished	Distinguished	Not Demonstrated (Comment Required)
Observation	**Element Ic. Teachers lead the teaching profession.** Teachers strive to improve the teaching profession. They contribute to the establishment of positive working conditions in their school. They actively participate in and advocate for decision-making structures in education and government that take advantage of the expertise of teachers. Teachers promote professional growth for all educators and collaborate with their colleagues to improve the profession.				
	❑ Has knowledge of opportunities and the need for professional growth and begins to establish relationships with colleagues.	. . . and Contributes to the: ❑ improvement of the profession through professional growth. ❑ establishment of positive working relationships. ❑ school's decision-making processes as required.	. . . and ❑ Promotes positive working relationships through professional growth activities and collaboration.	. . . and ❑ Seeks opportunities to lead professional growth activities and decision-making processes.	
	Element Id. Teachers advocate for schools and students. Teachers advocate for positive change in policies and practices affecting student learning. They participate in the implementation of initiatives to improve the education of students.				
	❑ Knows about the policies and practices affecting student learning.	. . . and ❑ Supports positive change in policies and practices affecting student learning.	. . . and ❑ Participates in developing policies and practices to improve student learning.	. . . and ❑ Actively participates, promotes, and provides strong supporting evidence for implementation of initiatives to improve education.	
	Element Ie. Teachers demonstrate high ethical standards. Teachers demonstrate ethical principles including honesty, integrity, fair treatment, and respect for others. Teachers uphold the *Code of Ethics for North Carolina Educators* (effective June 1, 1997) and the *Standards for Professional Conduct* adopted April 1, 1998. (www.ncptsc.org)				
	❑ Understands the importance of ethical behavior as outlined in the *Code of Ethics for North Carolina Educators* and the *Standards for Professional Conduct.*	. . . and ❑ Demonstrates ethical behavior through adherence to the *Code of Ethics for North Carolina Educators* and the *Standards for Professional Conduct.*	. . . and ❑ Knows and upholds the *Code of Ethics for North Carolina Educators* and the *Standards for Professional Conduct.*	. . . and ❑ Models the tenets of the *Code of Ethics for North Carolina Educators* and the *Standards for Professional Conduct* and encourages others to do the same.	

Comments:

Examples of Artifacts:

❑ Lesson plans	❑ Class rules and procedures	❑ National Board Certification
❑ Journals	❑ Participation in the Teacher Working	❑ Discipline Records
❑ Student handbooks	Condition Survey	❑ _____
❑ Student work	❑ Professional Learning Communities	❑ _____
❑ School improvement planning	❑ Membership in professional organizations	❑ _____
❑ Service on committees	❑ Formal and informal mentoring	❑ _____
❑ Relevant data	❑ Surveys	

23

Standard II: Teachers Establish a Respectful Environment for a Diverse Population of Students

Observation	**Element IIa. Teachers provide an environment in which each child has a positive, nurturing relationship with caring adults.** Teachers encourage an environment that is inviting, respectful, supportive, inclusive, and flexible.				
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated (Comment Required)**
✓	❑ Appreciates and understands the need to establish nurturing relationships.	. . . and ❑ Establishes an inviting, respectful, inclusive, flexible, and supportive learning environment.	. . . and ❑ Maintains a positive and nurturing learning environment.	. . . and ❑ Encourages and advises others to provide a nurturing and positive learning environment for all students.	
	Element IIb. Teachers embrace diversity in the school community and in the world. Teachers demonstrate their knowledge of the history of diverse cultures and their role in shaping global issues. They actively select materials and develop lessons that counteract stereotypes and incorporate histories and contributions of all cultures. Teachers recognize the influence of race, ethnicity, gender, religion, and other aspects of culture on a student's development and personality. Teachers strive to understand how a student's culture and background may influence his or her school performance. Teachers consider and incorporate different points of view in their instruction.					
✓	❑ Acknowledges that diverse cultures impact the world.	. . . and ❑ Displays knowledge of diverse cultures, their histories, and their roles in shaping global issues.	. . . and ❑ Uses materials or lessons that counteract stereotypes and acknowledges the contributions of all cultures.	. . . and ❑ Promotes a deep understanding of cultures through the integration of culturally sensitive materials and ideas throughout the curriculum.	
✓	❑ Demonstrates awareness of the diversity of students in the classroom.	❑ Acknowledges the influence of race, ethnicity, gender, religion, socio-economics, and culture on a student's development and attitudes.	❑ Consistently incorporates different points of view in instruction.	❑ Capitalizes on diversity as an asset in the classroom.	
	Element IIc. Teachers treat students as individuals. Teachers maintain high expectations, including graduation from high school, for students of all backgrounds. Teachers appreciate the differences and value the contributions of each student in the learning environment by building positive, appropriate relationships.					
✓	❑ Holds high expectations of students.	. . . and ❑ Communicates high expectations for all students.	. . . and ❑ Encourages and values contributions of students, regardless of background or ability.	. . . and ❑ Helps students hold high expectations for themselves and their peers.	

Observation	**Element IId. Teachers adapt their teaching for the benefit of students with special needs.** Teachers collaborate with the range of support specialists to help meet the special needs of all students. Through inclusion and other models of effective practice, teachers engage students to ensure that their needs are met.				
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated (Comment Required)**
✓	❑ Recognizes that students have a variety of learning needs.	. . . and ❑ Collaborates with specialists who can support the special learning needs of students.	. . . and ❑ Understands the roles of and collaborates with the full range of support specialists to help meet the special needs of all students.	. . . and ❑ Anticipates the unique learning needs of students and solicits assistance from within and outside the school to address those needs.	
✓	❑ Is knowledgeable of effective practices for students with special needs.	❑ Provides unique learning opportunities such as inclusion and research-based, effective practices for students with special needs.	❑ Effectively engages special needs students in learning activities and ensures their unique learning needs are met.	❑ Adapts instruction for the benefit of students with special needs and helps colleagues do the same for their students.	
	Element IIe. Teachers work collaboratively with the families and significant adults in the lives of their students. Teachers recognize that educating children is a shared responsibility involving the school, parents or guardians, and the community. Teachers improve communication and collaboration between the school and the home and community in order to promote trust and understanding and build partnerships with all segments of the school community. Teachers seek solutions to overcome cultural and economic obstacles that may stand in the way of effective family and community involvement in the education of their students.				
	❑ Responds to family and community concerns.	. . . and ❑ Communicates and collaborates with the home and community for the benefit of students.	. . . and ❑ Recognizes obstacles to family and community participation and conscientiously seeks solutions to overcome them.	. . . and ❑ Promotes trust and understanding throughout the school community.	

Comments:

Examples of Artifacts:

❑ Student profiles
❑ Student surveys
❑ Cooperation with ESL teachers
❑ Lessons that integrate international content
❑ Documentation of referral data and use of IEPs

❑ Communications w/parents/community
❑ Professional development on cultural attitudes and awareness
❑ Use of technology to incorporate cultural awareness into lessons

❑ _____
❑ _____
❑ _____
❑ _____
❑ _____

Standard III: Teachers Know the Content They Teach

Observation		Developing	Proficient	Accomplished	Distinguished	Not Demonstrated (Comment Required)
	Element IIIa. Teachers align their instruction with the *North Carolina Standard Course of Study*. In order to enhance the *North Carolina Standard Course of Study*, teachers investigate the content standards developed by professional organizations in their specialty area. They develop and apply strategies to make the curriculum rigorous and relevant for all students and provide a balanced curriculum that enhances literacy skills. Elementary teachers have explicit and thorough preparation in literacy instruction. Middle and high school teachers incorporate literacy instruction within the content area or discipline.					
✔		❑ Demonstrates an awareness of the *North Carolina Standard Course of Study* and references it in the preparation of lesson plans.	. . . and ❑ Understands the *North Carolina Standard Course of Study*, uses it in preparation of lesson plans, and applies strategies to make the curriculum rigorous and relevant.	. . . and ❑ Develops and applies strategies based on the *North Carolina Standard Course of Study* and standards developed by professional organizations to make the curriculum balanced, rigorous and relevant.	. . . and ❑ Assists colleagues in applying such strategies in their classrooms.	
✔		❑ *Elementary:* Begins to integrate literacy instruction in selected lessons.	❑ *Elementary:* Integrates effective literacy instruction throughout the curriculum.	❑ *Elementary:* Evaluates and reflects upon the effectiveness of literacy instruction.	❑ *Elementary:* Makes necessary changes to instructional practice to improve student learning.	
✔		❑ *Secondary:* Recognizes the importance of integrating literacy strategies within the content areas.	❑ *Secondary:* Incorporates a wide variety of literacy skills within content areas to enhance learning.	❑ *Secondary:* Evaluates and reflects upon the effectiveness of literacy instruction within content areas.	❑ *Secondary:* Makes necessary changes to instructional practice to improve student learning.	
	Element IIIb. Teachers know the content appropriate to their teaching specialty. Teachers bring a richness and depth of understanding to their classrooms by knowing their subjects beyond the content they are expected to teach and by directing students' natural curiosity into an interest in learning. Elementary teachers have broad knowledge across disciplines. Middle school and high school teachers have depth in one or more specific content areas or disciplines.					
✔		❑ Demonstrates a basic level of content knowledge in the teaching specialty to which assigned.	. . . and ❑ Demonstrates an appropriate level of content knowledge in the teaching specialty to which assigned.	. . . and ❑ Applies knowledge of subject beyond the content in assigned teaching specialty. Motivates students to investigate the content area to expand their knowledge and satisfy their natural curiosity.	. . . and ❑ Extends knowledge of subject beyond content in their teaching specialty and sparks students' curiosity for learning beyond the required course work.	

Observation	**Element IIIc. Teachers recognize the interconnectedness of content areas/disciplines.** Teachers know the links and vertical alignment of the grade or subject they teach and the *North Carolina Standard Course of Study*. Teachers understand how the content they teach relates to other disciplines in order to deepen understanding and connect learning for students. Teachers promote global awareness and its relevance to subjects they teach.				
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated (Comment Required)**
✓	❏ Understands the links between grade/subject and the *North Carolina Standard Course of Study*.	. . . and ❏ Demonstrates knowledge of links between grade/subject and the *North Carolina Standard Course of Study*.	. . . and ❏ Demonstrates knowledge of the links and vertical alignment of the grade or subject area and the *North Carolina Standard Course of Study*. Relates content to other disciplines.	. . . and ❏ Collaborates with teachers from other grades or subject areas to establish links between disciplines and influence school-wide curriculum and teaching practice.	
✓	❏ Displays global awareness.	❏ Promotes global awareness and its relevance to the subjects.	❏ Integrates global awareness activities throughout lesson plans and classroom instructional practices.	❏ Promotes global awareness and its relevance to all faculty members, influencing curriculum and teaching practices throughout the school.	
	Element IIId. Teachers make instruction relevant to students. Teachers incorporate 21st century life skills into their teaching deliberately, strategically, and broadly. These skills include leadership, ethics, accountability, adaptability, personal productivity, personal responsibility, people skills, self- direction, and social responsibility. Teachers help their students understand the relationship between the *North Carolina Standard Course of Study* and 21st century content, which includes global awareness; financial, economic, business and entrepreneurial literacy; civic literacy; and health awareness.				
✓	❏ Identifies relationships between the *North Carolina Standard Course of Study* and life in the 21st century.	. . . and ❏ Identifies relationships between the core content and 21st century content.	. . . and ❏ Integrates core content and 21st century content throughout lesson plans and classroom instructional practices.	. . . and ❏ Deepens students' understandings of 21st century skills and helps them make their own connections and develop new skills.	

Comments:

Examples of Artifacts:

❏ Display of creative student work ❏ Content standards ❏ _____

❏ Use of NC Standard Course of Study ❏ _____ ❏ _____

❏ Lesson plans ❏ _____ ❏ _____

Standard IV: Teachers facilitate learning for their students

Observation	Element IVa. Teachers know the ways in which learning takes place, and they know the appropriate levels of intellectual, physical, social, and emotional development of their students. Teachers know how students think and learn. Teachers understand the influences that affect individual student learning (development, culture, language proficiency, etc.) and differentiate their instruction accordingly. Teachers keep abreast of evolving research about student learning. They adapt resources to address the strengths and weaknesses of their students.				
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated** (Comment Required)
✔	❑ Understands developmental levels of students and recognizes the need to differentiate instruction.	. . . and ❑ Understands developmental levels of students and appropriately differentiates instruction.	. . . and ❑ Identifies appropriate developmental levels of students and consistently and appropriately differentiates instruction.	. . . and ❑ Encourages and guides colleagues to adapt instruction to align with students' developmental levels.	
✔		❑ Assesses resources needed to address strengths and weaknesses of students.	❑ Reviews and uses alternative resources or adapts existing resources to take advantage of student strengths or address weaknesses.	❑ Stays abreast of current research about student learning and emerging resources and encourages the school to adopt or adapt them for the benefit of all students.	

Observation	Element IVb. Teachers plan instruction appropriate for their students. Teachers collaborate with their colleagues and use a variety of data sources for short- and long-range planning based on the *North Carolina Standard Course of Study*. These plans reflect an understanding of how students learn. Teachers engage students in the learning process. They understand that instructional plans must be consistently monitored and modified to enhance learning. Teachers make the curriculum responsive to cultural differences and individual learning needs.				
✔	❑ Recognizes data sources important to planning instruction.	. . . and ❑ Uses a variety of data for short- and long-range planning of instruction. Monitors and modifies instructional plans to enhance student learning.	. . . and ❑ Monitors student performance and responds to individual learning needs in order to engage students in learning.	. . . and ❑ Monitors student performance and responds to cultural diversity and learning needs through the school improvement process.	

Observation	Element IVc Teachers use a variety of instructional methods. Teachers choose the methods and techniques that are most effective in meeting the needs of their students as they strive to eliminate achievement gaps. Teachers employ a wide range of techniques including information and communication technology, learning styles, and differentiated instruction.				
✔	❑ Demonstrates awareness of the variety of methods and materials necessary to meet the needs of all students.	. . . and ❑ Demonstrates awareness or use of appropriate methods and materials necessary to meet the needs of all students.	. . . and ❑ Ensures the success of all students through the selection and utilization of appropriate methods and materials.	. . . and ❑ Stays abreast of emerging research areas and new and innovative materials and incorporates them into lesson plans and instructional strategies.	

Observation	Element IVd. Teachers integrate and utilize technology in their instruction. Teachers know when and how to use technology to maximize student learning. Teachers help students use technology to learn content, think critically, solve problems, discern reliability, use information, communicate, innovate, and collaborate.				
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated (Comment Required)**
✔	❑ Assesses effective types of technology to use for instruction.	. . . and ❑ Demonstrates knowledge of how to utilize technology in instruction.	. . . and ❑ Integrates technology with instruction to maximize student learning.	. . . and ❑ Provides evidence of student engagement in higher level thinking skills through the integration of technology.	
	Element IVe. Teachers help students develop critical-thinking and problem-solving skills. Teachers encourage students to ask questions, think creatively, develop and test innovative ideas, synthesize knowledge, and draw conclusions. They help students exercise and communicate sound reasoning; understand connections; make complex choices; and frame, analyze, and solve problems.				
✔	❑ Understands the importance of developing students' critical thinking and problem solving skills.	. . . and ❑ Demonstrates knowledge of processes needed to support students in acquiring critical thinking skills and problem-solving skills.	. . . and Teaches students the processes needed to: ❑ think creatively and critically, ❑ develop and test innovative ideas, ❑ synthesize knowledge, ❑ draw conclusions, ❑ exercise and communicate sound reasoning, ❑ understand connections, ❑ make complex choices, and ❑ frame, analyze and solve problems.	. . . and ❑ Encourages and assists teachers throughout the school to integrate critical thinking and problem solving skills into their instructional practices.	
	Element IVf. Teachers help students work in teams and develop leadership qualities. Teachers teach the importance of cooperation and collaboration. They organize learning teams in order to help students define roles, strengthen social ties, improve communication and collaborative skills, interact with people from different cultures and backgrounds, and develop leadership qualities.				
✔	❑ Provides opportunities for cooperation, collaboration, and leadership through student learning teams.	. . . and ❑ Organizes student learning teams for the purpose of developing cooperation, collaboration, and student leadership.	. . . and ❑ Encourages students to create and manage learning teams.	. . . and ❑ Fosters the development of student leadership and teamwork skills to be used beyond the classroom.	

Element IVg. Teachers communicate effectively. Teachers communicate in ways that are clearly understood by their students. They are perceptive listeners and are able to communicate with students in a variety of ways even when language is a barrier. Teachers help students articulate thoughts and ideas clearly and effectively.

Observation	Developing	Proficient	Accomplished	Distinguished	Not Demonstrated (Comment Required)
		. . . and	. . . and	. . . and	
✔	❑ Demonstrates the ability to effectively communicate with students.	❑ Uses a variety of methods for communication with all students.	❑ Creates a variety of methods to communicate with all students.	❑ Anticipates possible student misunderstandings and proactively develops teaching techniques to mitigate concerns.	
✔	❑ Provides opportunities for students to articulate thoughts and ideas.	❑ Consistently encourages and supports students to articulate thoughts and ideas clearly and effectively.	❑ Establishes classroom practices which encourage all students to develop effective communication skills.	❑ Establishes school-wide and grade appropriate vehicles to encourage students throughout the school to develop effective communication skills.	

Element IVh. Teachers use a variety of methods to assess what each student has learned. Teachers use multiple indicators, including formative and summative assessments, to evaluate student progress and growth as they strive to eliminate achievement gaps. Teachers provide opportunities, methods, feedback, and tools for students to assess themselves and each other. Teachers use 21st century assessment systems to inform instruction and demonstrate evidence of students' 21st century knowledge, skills, performance, and dispositions.

Observation	Developing	Proficient	Accomplished	Distinguished	Not Demonstrated (Comment Required)
		. . . and	. . . and	. . . and	
✔	❑ Uses indicators to monitor and evaluate student progress.	❑ Uses multiple indicators, both formative and summative, to monitor and evaluate student progress and to inform instruction.	❑ Uses the information gained from the assessment activities to improve teaching practice and student learning.	❑ Teaches students and encourages them to use peer and self-assessment feedback to assess their own learning.	
✔	❑ Assesses students in the attainment of 21st century knowledge, skills, and dispositions.	❑ Provides evidence that students attain 21st century knowledge, skills and dispositions.	❑ Provides opportunities for students to assess themselves and others.	❑ Encourages and guides colleagues to assess 21st century skills, knowledge, and dispositions and to use the assessment information to adjust their instructional practice.	

Comments:

Examples of Artifacts:

❑ Lesson plans	❑ Documentation of differentiated instruction	❑ _____
❑ Display of technology used	❑ Materials used to promote critical thinking	❑ _____
❑ Professional development	and problem solving	❑ _____
❑ Use of student learning teams	❑ Collaborative lesson planning	❑ _____

Standard V: Teachers Reflect on Their Practice

Observation					
Element Va. Teachers analyze student learning. Teachers think systematically and critically about student learning in their classrooms and schools: why learning happens and what can be done to improve achievement. Teachers collect and analyze student performance data to improve school and classroom effectiveness. They adapt their practice based on research and data to best meet the needs of students.					
	Developing	**Proficient**	**Accomplished**	**Distinguished**	**Not Demonstrated (Comment Required)**
	❑ Recognizes the need to improve student learning in the classroom.	. . . and ❑ Provides ideas about what can be done to improve student learning in their classroom.	. . . and ❑ Thinks systematically and critically about learning in their classroom: why learning happens and what can be done to improve student achievement.	. . . and ❑ Provides a detailed analysis about what can be done to improve student learning and uses such analyses to adapt instructional practices and materials within the classroom and at the school level.	
Element Vb. Teachers link professional growth to their professional goals. Teachers participate in continued, high-quality professional development that reflects a global view of educational practices; includes 21st century skills and knowledge; aligns with the State Board of Education priorities; and meets the needs of students and their own professional growth.					
	❑ Understands the importance of professional development.	. . . and ❑ Participates in professional development aligned with professional goals.	. . . and ❑ Participates in professional development activities aligned with goals and student needs.	. . . and ❑ Applies and implements knowledge and skills attained from professional development consistent with its intent.	
Element Vc. Teachers function effectively in a complex, dynamic environment. Understanding that change is constant, teachers actively investigate and consider new ideas that improve teaching and learning. They adapt their practice based on research and data to best meet the needs of their students.					
	❑ Is knowledgeable of current research-based approaches to teaching and learning.	. . . and ❑ Considers and uses a variety of research-based approaches to improve teaching and learning.	. . . and ❑ Actively investigates and considers alternative research-based approaches to improve teaching and learning and uses such approaches appropriately.	. . . and ❑ Adapts professional practice based on data and evaluates impact on student learning.	

Comments:

Examples of Artifacts:

❑ Lesson plans

❑ Formative assessments

❑ Student work

❑ Professional Development Plan

❑ Completion of professional development

❑ Participation in professional learning community

❑ Formative and summative assessment data

❑ _____

❑ _____

❑ _____

❑ _____

REFERENCES

Aaronson, D., Barrow, L., & Sanders, W. (2007, January). Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics, 25*(1), 95-135. Retrieved from http://www.jstor.org/stable/10.1086/508733

Alicias, E. R. (2005). Toward an objective evaluation of teacher performance: The use of variance partitioning analysis, VPA. *Education Policy Analysis Archives, 13*(30), 1-13. doi: http://dx.doi.org/10.14507/epaa.v13n30.2005

Allison, P.D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.

American Recovery and Reinvestment Act of 2009. (2009). In GovTrack.us (database of federal legislation). Retrieved November 10, 2012, from http://www.govtrack.us/congress/bills/111/hr1

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*(2), 65-75. Retrieved from http://www.jstor.org/stable/30137966

Amrein-Beardsley, A. (2009). Value-added tests: Buyer, be aware. *Educational Leadership, 67*(3), 38-42. Retrieved from http://www.ascd.org/publications/educational_leadership/nov09/vol67/num03/Value-Added_Tests@_Buyers,_Be_Aware.aspx

Amrein-Beardsley, A., & Collins, C. (2012). The SAS Education Value-Added Assessment System (SAS EVAAS) in the Houston Independent School District (HISD): Intended and unintended consequences. *Education Policy Analysis Archives, 20*(12). Retrieved from http://epaa.asu.edu/ojs/article/view/1096

Amrein-Beardsley, A., Collins, C., Polasky, S.A., & Sloat, E.F. (2013). Value-Added Model (VAM) research for educational policy: Framing the issue. *Education Policy Analysis Archives, 21*(4). Retrieved from http://epaa.asu.edu/ojs/article/view/1311/1350

André-Bechely, L. (2005). Public school choice at the intersection of voluntary integration and not-so-good neighborhood schools: Lessons from parents' experiences*. Educational* Administration Quarterly, 41(2), 267-305

Andrejko, L. (2004). Value-added assessment: a view from a practitioner. *Journal of Educational and Behavioral Statistics, 29*(1), 7-9. Retrieved from http://www.jstor.org/stable/3701304

Ansalone, G. (2006). Tracking: A return to jim crow. *Race, Gender & Class, 13*(1), 144-153.

Appropriations Act of 2014, N.C. Gen. Stat. § 115C-472.16 (2014).

ArcGIS (2016). NCES urban-centric locale codes for 2013-2014 school year as collected by the National Center for Education Statistics (NCES). Retrieved from https://www.arcgis.com/home/item.html?id=e50e883de26649a4a2981e3455e384a5

Archbald, D., & Farley-Ripple, E.N. (2012). Predictors of placement in lower level versus higher level high school mathematics. *High School Journal, 96*(1), 33–51. doi:10.1353/hsj.2012.0014

Ariz. Rev. Stat. Ann. § 15-977 (2007)

Armor, D., Conry-Oseguera, P., Fox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, CA: RAND Corporation.

Au, W. (2010, Winter). Neither fair nor accurate: Research-based reasons why high-stakes tests should not be used to evaluate teachers. *Rethinking schools.* Retrieved from http://www.rethinkingschools.org/archive/25_02/25_02_au.shtml

Bagdon, E. M. (1985). Teacher competency testing and merit pay proposals under the equal protection clause and title VII. *Washington University Journal of Urban and Contemporary Law, 28*, 251-272. Retrieved from http://www.heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/waucl28&start_page=251&id=259

Ball, J. (2015, March 2). Which NC teachers should be paid more? *Citizen Times*. Retrieved from http://www.citizen-times.com/story/news/local/2015/03/01/nc-teachers-paid/24234023/

Ballou, D. (2002). Sizing up test scores. *Education Next, 2*(2). Retrieved from http://libproxy.lib.unc.edu/login?url=http://search.proquest.com.libproxy.lib.unc.edu/docview/1237764582?accountid=14244

Ballou, D. (2012). *Review of "The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood"*. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/thinktank/review-long-term-impacts

Ballou, D., Sanders, W., & Wright, P. (2004, Spring). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*(1), 37-65. Retrieved from http://www.jstor.org/stable/3701306

Baltagi, B. H., (2005). *Econometric Analysis of Panel Data*. Chichester, England: John Wiley & Sons.

Banchero, S., & Rutland, M. (2013, July 26). North Carolina ends pay boosts for teacher master's degrees. *Wall Street Jounral*. Retrieved from https://www.wsj.com/articles/SB10001424127887323971204578630312785220612

Bastian, K.C., Henry, G.T., & Thompson, C.L. (2013, Fall). Incorporating access to more effective teachers into assessments of educational resource equity. *Education Finance and Policy, 8*(4), 560-580. doi: 10.1162/EDFP_a_00113.

Battey, D. (2013). Access to mathematics: A possessive investment in whiteness. *Curriculum Inquiry, 43*(3), 332–359. doi:10.1111/curi.12015

Bird, R.B., & Smith. A. (2005). Signaling theory, strategic interaction, and symbolic capital. *Current Anthropology, 46*(2), 221-248. Retrieved from http://www.jstor.org/stable/10.1086/427115

Bill & Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching Project*. MET Project Research Paper. Seattle, Washington: Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The effectiveness and retention of teachers with prior career experience. *Economics of Education Review, 30*(6), 1229-1241. Retrieved from https://doi.org/10.1016/j.econedurev.2011.08.004.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management, (27)*4, 793-818. Retrieved from http://www.jstor.org/stable/30163563

Brodsky, A., DeCesare, D., & Kramer-Wine, J. (2010). Design and implementation considerations for alternative teacher compensation programs. *Theory Into Practice*, *49*, 213-222. doi:10.1080/00405841

Brophy, J. (1973). Stability of teacher effectiveness. *American Educational Research Journal, 10*(3), 245-252. doi:10.2307/1161888.

Brown, K.D. (2014). Teaching in color: A critical race theory education analysis of the literature on preservice teachers of color and teacher education in the US. *Race Ethnicity and Education, 17*(3), 326-345. http://dx.doi.org.libproxy.lib.unc.edu/10.1080/13613324.2013.832921

Buddin, R. (2011). *Measuring teacher and school effectiveness at improving student achievement in Los Angeles elementary schools*. (Working Paper No. 31963). Retrieved from http://mpra.ub.uni-muenchen.de/31963

Busentitz, L.W., Fiet, J.O., & Moesel, D.D. (2005). Signaling in venture capitalist-new venture team funding decisions: Does it indicate long-term venture outcomes? *Entrepreneurship Theory and Practice, 29*(1), 1-12. doi:0.1111/j.1540-6520.2005.00066.x

Calhoun, F.S, & Protheroe, N.J. (1983). Merit pay plans for teachers: status & descriptions 3. Arlington, VA: Educational Research Service. (ED236771)

Cameron.A.C., & Miller, D.L. (2015). A practitioner's guide to cluster-robust inference. *The Journal of of Human Resources (50)*2, 317-372. Retrieved from https://muse.jhu.edu/article/581178

Cantrell, S., & Kane, T. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. MET Project Research Paper. Seattle, Washington: Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf

Casper, W.J., Wayne, J.H., & Manegold, J.G. (2013). Who will we recruit? Targeting deep and surface-level diversity with human resource policy advertising. *Human Resource Management, 52*(3), 311-332. doi:10.1002/hrm.21530.

Certo, S.T. (2003). Influencing initial public offering investors with prestige: Signaling with board structures. *The Academy of Management Review, 28*(3), 432-446. Retrieved from http://www.jstor.org/stable/30040731

Chetty, R., Friedman, J.N., & Rockoff, J.E. (2011). The long-term impacts of teachers: teacher value-added and student outcomes in adulthood. *National Bureau for Economic Research* (Working Paper No. 17699). Cambridge, MA: NEBR. Retrieved from http://www.nber.org/papers/w17699

Chingos, M.M., & Peterson, P.E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review, 30*(3), 449-465. doi: doi:10.1016/j.econedurev.2010.12.010

Chingos, M.M., & West, M.R. (2012). Do more effective teachers earn more outside of the classroom? *Education Finance and Policy, 7*(1), 8-43. doi:10.1162/EDFP_a_00052

Citron, C. H. (1985). An overview of legal issues in teacher quality. *Journal of Law & Education, 14*(3), 277-308. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/jle14&id=287

Clotfelter, T. Ladd, H.F., & Vigor, J.L. (2006, Fall). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of Human Resources, 41*(4), 778-820. Retrieved from http://www.jstor.org/stable/40057291

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007a). How and why do teacher credentials matter for student achievement? *National Center for Analysis of Longitudinal Data in Education Research* (Working Paper No. 11). Washington, DC: Urban Institute. Retrieved from http://www.nber.org/papers/w12828

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007b). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*, 673-682. doi:10.1016/j.econedurev.2007.10.002

Cody, C.A., McFarland, J.., Moore, J.E., & Preston, J. (2010). *The evolution of growth models*. Public Schools of North Carolina. Retrieved from http://www.dpi.state.nc.us/docs/intern-research/reports/growth.pdf

Coleman, J. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.

Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS Education Value-Added Assessment System (EVAAS). *Education Policy Analysis Archives, 22*(98). Retrieved from http://dx.doi.org/10.14507/epaa.v22.1594

Collins, C. & Amrein-Beardsley, A. (2014). Putting growth and value-added models on the map: A national overview. *Teachers College Record, 16*(1). Retrieved from http://www.tcrecord.org/Content.asp?ContentID=17291

Colo. Rev. Stat. Ann. § 22-63-401 (2014)

Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1). doi:0.1177/0149206310388419

Cornbleth, C. (2006). Curriculum and students: Diverting the public interest. In G. Ladson-Billings & W. F. Tate (Eds.), *Education research in the public interest* (pp. 199-212). New York, NY: Teachers College Press.

Costello, R., Elson, P., & Schacter, J. (2008). An introduction to value-added analysis. *Catholic Education, 12*(2), 193-203. Retrieved from https://auth-lib-unc-edu.libproxy.lib.unc.edu/ezproxy_auth.php?url=http://search.ebscohost.com.libproxy.lib.unc.edu/login.aspx?direct=true&db=eft&AN=508022284&site=ehost-live&scope=site

Creswell, J. W. (2012). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson.

Croninger, R.G, Rice, J.K., Rathbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review, 26*(3), 312-324. doi:10.1016/j.econedurev.2005.05.008

Daniel, P. T. K., & Dyson, M.R. (2009-2010). Bringing every child forward: Lessons learned under no child left behind & (and) a roadmap for Obama's educational reform. *Temple Political & Civil Rights Law Review, 19*, 63-90. Retrieved from: http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/tempcr19&id=65&collection=journals&index

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan, 93*(6), 8-15. Retrieved from http://www.jstor.org/stable/41497541

David, J. L. (2010). Using value-added measures to evaluate teachers. *Educational Leadership, 67*(8), 81-82. Retrieved from https://auth-lib-unc-edu.libproxy.lib.unc.edu/ezproxy_auth.php?url=http://search.ebscohost.com.libproxy.lib.unc.edu/login.aspx?direct=true&db=aph&AN=50303085&site=ehost-live&scope=site

Dee, T.S., & Keys, B.J. (2004, Summer). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management, 23*(3), 471-488. Retrieved from http://www.jstor.org/stable/3326262

Dee, T. (2005). A Teacher like Me: Does Race, Ethnicity, or Gender Matter? *The American Economic Review, 95*(2), 158-165. Retrieved from http://www.jstor.org/stable/4132809

Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York, NY: New Press.

Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, H.R. 310, 109th Cong., 1st Sess. (2006). Retrieved from http://www.govtrack.us/congress/bills/109/hr3010/text

Dewey, J. Husted, T.A., & Kenney, L.W. (2000). The ineffectiveness of School Inputs: A product misspecification? *Economics of Education Review, 19*(1), 27-45. doi:10.1016/S0272-7757(99)00015-1

Dewitt, D. (2014, April 9). Ranking teachers: NC bets big on a complicated stats model. *North Carolina Public Radio – WUNC*. Retrieved from http://wunc.org/post/ranking-teachers-nc-bets-big-complicated-stats-model

Edmonds, T. D. (2012). Contracting away success: The way teacher collective bargaining agreements are undermining the education of America's children. *Columbia Journal of Race and Law, 2*, 199-240. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/cjoral2&id=199

Education Policy Initiative at Carolina (2014). EPIC. Retrieved from http://publicpolicy.unc.edu/epic-2/

Ehrenber, R.G., & Brewer, D.J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review, 13*(1), 1-17. doi:10.1016/0272-7757(94)90019-1

Ehrhart, K.H., & Ziegert, J.C. (2005). Why are individuals attracted to organizations? *Journal of Management, 31*(6), 901-919. doi:10.1177/0149206305279759.

Epstein, K. K. (2005). The whitening of the American teaching force: A problem of recruitment or a problem of racism? *Social Justice, 32(3)*, 89-102. Retrieved from http://libproxy.lib.unc.edu/login?url=https://search-proquest-com.libproxy.lib.unc.edu/docview/231894917?accountid=14244

Feldman, S. (2000). 220,000 teachers a year: Putting first-class educators in every classroom public address. *Georgetown Journal on Poverty Law and Policy, 7*, 173-180. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/geojpovlp7&collection=journals&page=173

Figlio, D. N., & Kenny, L. W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics, 91*(5–6), 901-914. doi: 10.1016/j.jpubeco.2006.10.00

Filatotchev, I., & Bishop, K. (2002). Board composition, share ownership, and 'underpricing' of U.K. IPO firms. *Strategic Management Journal, 23*(10), 941-955. Retrieved from http://www.jstor.org/stable/3094415

Fortune, J.C., & O'Neil, J.S. (1994). Production function analyses and the study of educational funding equity: A methodological critique. *Journal of Education Finance, 20*(1)*, 21-46. Retrieved from http://www.jstor.org/stable/40703911

Gay, L.R., Mills, G.E., & Airasian, P. (2006). *Educational research: Competencies for Analysis and Applications.* Upper Saddle River, NJ: Pearson Education.

Goldhaber, D. (2002, Spring). The mystery of good teaching. *Education Next*, 2(1), 50-55. Retrieved from http://libproxy.lib.unc.edu/login?url=http://search.proquest.com.libproxy.lib.unc.edu/docview/1237764529?accountid=14244

Goldhaber, D., & Brewer, D. (1997, Summer). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources, 32*(2), 505-520. Retrieved from http://www.jstor.org/stable/146181

Goldhaber, D., & Brewer, D. (1998). When should we reward degrees for teachers? *Phi Delta Kappan,* 80(2), 134-138. Retrieved from http://www.jstor.org/stable/20439382

Goldhaber, D., DeArmond, M., & DeBurgomaster, S. (2011). Teacher attitudes about compensation reform: Implications for reform implementation. *Industrial and Labor Relations Review, 64*(3). Retrieved from http://www.jstor.org/stable/41149474

Goldhaber, D., & Hansen, M. (2010). Race, gender, and teacher testing: How objective a tool is teacher licensure testing? *American Educational Research Journal 47*, 218-251. Retrieved from http://www.jstor.org/stable/40645423

Goldrick, L. (2002). *Improving teacher evaluation to improve teaching quality*. Washington, DC: National Governors Association. Retrieved from http://www.nga.org/cda/files/1202IMPROVINGTEACHEVAL.pdf

Goodman, M. R., & Melia, P. K. (1988-1989). Search for excellence in the nation's schools: Defining a new policy agenda, the public administration comments. *Hamline Journal of Public Law and Policy,9*, 185-208. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/hplp9&id=191

Greenwald, R., Hedges, L.V., & Laine, R.D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*(3), 361-396. Retrieved from http://www.jstor.org/stable/1170528

Hanushek, E.A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal Economic Literature, 24*(3), 1141-1177. Retrieved from http://www.jstor.org/stable/2725865

Hanushek, E.A. (1989). The impact of differential expenditures on school performance. *Educational Researcher, 18*(4), 45-51. Retrieved from http://www.jstor.org/stable/1176650

Hanushek, E.A. (1991). When school finance reform may not be good policy. *Harvard Journal on Legislation, 28*(2), 424-456. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/hjl28&id=429

Hanushek, E.A. (1997). Assessing the effects of school resource on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141-164. Retrieved from http://www.jstor.org/stable/1164207

Hanushek, E.A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479. Retrieved from doi:10.1016/j.econedurev.2010.12.006

Hanushek, E.A., Kain, J.F., O'Brien, D.M., & Rivkin, S.G. (2005). The market for teacher quality (Working Paper No. 11154). *National Bureau for Economic* Research. Retrieved from http://www.nber.org/papers/w11154

Hanushek, E.A. & Rivkin, S.G. (2006). Teacher quality. In E. Hanushek & F. Welsh (Eds.), *Handbook of the Economics of Education* (1051 – 1078). Welsh, Amsterdam, NL: Elsevier B.V.

Hanushek, E.A., & Rivkin, S.G. (2007, Spring). Pay, working conditions and teacher quality. *Future of Children, 17(*1), 69-86. Retrieved from http://www.jstor.org/stable/4150020

Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know*. Cambridge, Mass.: Harvard Education Press.

Harris, D.N., & Sass, T.R. (2008). *What makes for a good teacher and who can tell?* Florida State University, Department of Economics.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7–8), 798-812. doi: 10.1016/j.jpubeco.2010.11.009

Harrison, C., & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narratives surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education, 87*(5), 517-534. doi: 10.1080/0161956X.2012.723489

Hazi, H. M., & Garman, N. B. (1988). Teachers ask: Is there life after Madeline Hunter? *Phi Delta Kappan, 69*(9), 669–672. Retrieved from http://www.jstor.org/stable/20403745

Hazi, H.M., & Rucinski, D.A. (2009). Teacher evaluation as a policy target for improved student learning: a fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives, 17*(5), 1-22. Retrieved from http://epaa.asu.edu/epaa/v17n5/

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). An exchange: Part I: Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher, 23*(3), 5-14. Retrieved from http://www.jstor.org/stable/1177220

Henry, G.T., Fortner, C.K., & Bastian, K.C. (2012). The effects of experience and attrition for novice high-school science and mathematics teachers. *Science, 335*(6072), 1118-1121. doi:10.1126/science.1215343

Highhouse, S., Thornbury, E.E., & Little, I.S. (2007). Social-identity functions of attraction to organizations. *Organizational Behavior and Human Decision Processes, 103,* 134-146. doi: 10.1016/j.obhdp.2006.01.001.

Hill, J. (2011). Value-added history-social science effectiveness. *Social Studies Review, 50*(1), 20-24. Retrieved from http://libproxy.lib.unc.edu/login?url=http://search.proquest.com.libproxy.lib.unc.edu/docview/890178302?accountid=14244

Hilliard, A. G., III (1990). Back to Binet: The case against the use of IQ tests in the schools. *Contemporary Education, 61,* 184-189

Hochwarer, W.A., Ferris, G.R., Zinko, R., & Arnell, B. (2007). Reputation as a moderator of political behavior-work outcome relationships: A study investigation with convergent results. *Journal of Applied Psychology, 92*(2), 567-576. doi:10.1037/0021-9010.92.2.567

Ingle, W. K. (2009). Teacher Quality and Attrition in a US School District. *Journal Of Educational Administration*, *47*(5), 557-585. Retrieved from http://dx.doi.org.libproxy.lib.unc.edu/10.1108/09578230910981062

Ishii, J., & Rivkin, S.G. (2009). Impediments to the estimation of teacher value added. *Education Finance and Policy, 4*(4), 520-536. doi: 10.1162/edfp.2009.4.4.520.

Jacob, B., & Lefgren, L. (2005). Principals as agents: Subjective performance measurement in education. *National Bureau for Economic Research* (Working Paper No. 11463). Cambridge, MA: NEBR. Retrieved from http://www.nber.org/papers/w11463

Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics, 26*(1), 101-136. Retrieved from http://www.jstor.org/stable/10.1086/522974

Jensen, M.C., & Meckling, W.H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4)*, 305-360. doi:10.1016/0304-405X(76)90026-X

Kahne, J. (1994). Democratic communities, equity, and excellence: A Deweyan reframing of educational policy analysis. *Educational Evaluation and Policy Analysis, 16*(3), 233-248.

Kane, T.J., Rockoff, J.E., & Staiger, D. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*(6), 615-631. doi:10.1016/j.econedurev.2007.05.005

Kane, T.J., & Staiger, D.O. (2001). Improving school accountability measures (Working Paper No. 8156). *National Bureau for Economic Research* (Working Paper No. 11463). Cambridge, MA: NEBR. doi: 10.3386/w11154

Kane, T.J., & Staiger, D.O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives, 16*(4)*, 91-114. Retrieved from http://www.jstor.org/stable/3216916

Kane, T.J., & Staiger, D.O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research* (Working Paper 14607). Cambridge, MA: NEBR. Retrieved from http://www.nber.org/papers/w14607

Kane, T.J, & Staiger, D.O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* MET Project Research Paper. Seattle, Washington: Bill & Melinda Gates Foundation. Retrieved from http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/MET_Gathering_Feedback_Research_Paper.pdf

Kelley, C. (1997, Spring). Teacher compensation and organization. *Educational Evaluation and Policy Analysis, 19*(1), 15-28. Retrieved from http://www.jstor.org/stable/1164517

Kelly, S. (2007). The contours of tracking in North Carolina. *The High School Journal, 90*(4), pp. 15-31.

Kennedy, M.M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher, 39*(8), 591-598. doi:10.3102/00131189X10390804

Knoeppel, R.C., & Rinehart, J.S. (2008). Explaining the relationship between resources and student achievement: A methodological comparison of production function analysis and canonical analysis. *Educational Considerations, 35*(2), 22-40. Retrieved from https://auth-lib-unc-edu.libproxy.lib.unc.edu/ezproxy_auth.php?url=http://search.ebscohost.com.libproxy.lib.unc.edu/login.aspx?direct=true&db=eft&AN=508060946&site=ehost-live&scope=site

Koedell, C., & Betts (2007). Re-examining the role of teacher quality and the educational production function (Working Paper). Retrieved from Vanderbilt National Center on Performance Incentives website: https://my.vanderbilt.edu/performanceincentives/ncpi-publications/supplemental-studies/value-added-studies/re-examining-the-role-of-teacher-quality-and-the-educational-production-function/

Kupermintz, H. (2003). Teacher effects and effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis, 25*(3), 287-298. Retrieved from http://www.jstor.org/stable/3699496

Krieg, J. W. (2006). Teacher quality and attrition. *Economics of Education Review*, *25*, 13-27. Retrieved from https://doi.org/10.1016/j.econedurev.2004.09.004

Ladson-Billings, G. (2006, Winter). Teacher quality: Conversations on quality. *Rethinking Schools, 20*(2), 40. Retrieved from http://rethinkingschools.aidcvt.com/static/special_reports/quality_teachers/KeepingTeachersReport.pdf

Lample, J., & Shamsie, J. (2000). Critical push: Strategies for creating momentum in the motion picture industry. *Journal of Management, 26*(2), 233-257. doi:10.1016/S0149-2063(99)00044-6

Lefkowitz, J. (2000). The role of interpersonal affective regard in supervisory performance ratings: a literature review and proposed causal model. *Journal of Occupational and Organizational Psychology*, *73*(1), 67-85. Retrieved from http://search.proquest.com/docview/1634113769?accountid=14244

Lester, R.H., Certo, S.T., Dalton, C.M., Dalton, D.R., & Cannella, A.A. (2006). Initial public offering investor valuations: An examination of top management team prestige and environmental uncertainty. *Journal of Small Business Management, 44*(1), 1-26. Retrieved from Retrieved from http://search.proquest.com/docview/220959457?accountid=14244

Liang, G., & Akiba, M. (2011). Performance-related pay: District and teacher characteristics. *Journal of School Leadership, 21*(6), 845-870. Retrieved from http://go.galegroup.com.libproxy.lib.unc.edu/ps/i.do?id=GALE%7CA290622759&v=2.1&u=unc_main&it=r&p=AONE&sw=w

Lievens, F., & Highhouse, S. (2003, Spring). The relation of instrumental and symbolic attributes to a company's attractiveness as an employer. *Personnel Psychology, 56*(1), 75-201. Retrieved from http://search.proquest.com/docview/220143369?accountid=14244

Linn, R.L. (2008). Methodological issues in achieving school accountability. *Journal of Curriculum Studies, 40*, 699-711. doi:10.1080/00220270802105729

Linn, R.L., & Huag, C. (2002, Spring). Stability of school-building accountability scores and gains. *Education and Policy Analysis, 24*(1), 29-36. Retrieved from http://www.jstor.org/stable/3594145

Lockwood, J.R., McCaffrey, D.F., & Sass, T.R. (2008). *The intertemporal stability, of teacher effect estimates*. Paper presented at the National Conference on Value-Added Modeling. Sponsored by the Wisconsin Center for Education Research (WCER), Madison, WI. Retrieved from http://www.wcer.wisc.edu/news/events/VAM%20Conference%20Final%20Papers/IntertemporalStability_McCaffreySassLockwood.pdf

Loup, K.S., Garland, J.S., Ellett, C.D., & Rugutt, J.K. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practiced in our 100 largest districts. *Journal of Personnel Evaluation in Education, 10*(3)*,* 203-226.

Malin, M. H. (2009). Paradox of public sector labor law. *Indiana Law Journal, 84*(4), 1369-1400. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/indana84&id=1377

Mangiante, E.M.S. (2011). Teachers matter: Measures of teacher effectiveness in low-income minority schools. *Educational Assessment, Evaluation and Accountability, 23*(1), 41-63. Retrieved from http://dx.doi.org/10.1007/s11092-010-9107-x

Marshall, C. (1989). In Mitchell D. E., Wirt F. M. (Eds.), *Culture and education policy in the American states*. New York: Falmer Press.

McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., & Hamilton, L.S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: The Rand Corporation.

Miller, T., & Triana, M. (2009). Demographic diversity in the boardroom: Mediators of the board diversity-firm performance relationship. *Journal of Management Studies, 46*(5), 755-786. doi: 10.1111/j.1467-6486.2009.00839.x

Minn. Stat. § 122A.60 (2013)

Misco, T. (2008). Was that a result of my teaching? A brief exploration of value-added assessment. *The Clearing House, 82*(1), 11-14. doi:10.3200/TCHS.82.1.11-14

Miss. Code Ann. § 37-19-7 (2013)

Monk, D.H. (1989). The education production function: Its evolving role in policy analysis. *Educational Evaluation and Policy Analysis, 11*(1), 31-45. Retrieved from http://www.jstor.org/stable/1163714

Murnane, R.J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger Press.

Murnane, R.J., & Phillips, B. (1981). What do effective teachers of inner-city children have in common? *Social Science Research, 10*(1), 83-100. doi:10.1016/0049-089X(81)90007-7

National Commission on Excellence in Education. (1983, April). A nation at risk: The imperative for education reform.

National Council on Teacher Quality. (2011). State of the states: Trends and early lessons on teacher evaluations and effectiveness policies. New York. Retrieved from http://www.nctq.org/dmsView/State_of_the_States_Teacher_Evaluation_and_Effectiveness_Policies_NCTQ_Report

NC Department of Public Instruction (2015, July). 2015-2016 salary schedules. Retrieved from http://www.ncpublicschools.org/docs/fbs/finance/salary/schedules/2015-16schedules.pdf

NC Department of Public Instruction (2016). Technical documentation for 2016 EVAAS Analyses. Retrieved from https://ncdpi.sas.com/support/EVAAS-NC-TechnicalDocumentation-2016.pdf

NC Department of Public Instruction (2017). Title 1, Part A. Retrieved from http://www.ncpublicschools.org/program-monitoring/titleIA/

NC State Board of Education (2012, September). North Carolina Teacher Evaluation Process. Retrieved from http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237-257. Retrieved from http://www.jstor.org/stable/3699577

Oates, G.L. (2003). Teacher-student racial congruence, teacher perceptions, and test performance. *Social Science Quarterly, 84*(3), 508-525. doi: 10.1111/1540-6237.8403002

Omps, L. K. (2011-2012). Holding teachers accountable and rewarding those who perform: Evaluating a performance-based pay system for West Virginia student works. *West Virginia Law Review,114*(3), 1053-1088. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/wvb114&id=1063

Papay, J.P. (2010). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal, 48*(1), 163-193. Retrieved from http://www.jstor.org/stable/27975284

Peterson, P.E., Lastra-Anadon, C.X., Hanushek, E.A., & Woessmann, L. (2011). Are U.S. students ready to compete? *Education Next*, *11*(4). Retrieved from http://educationnext.org/are-u-s-students-ready-to-compete/

Podgursky, M.G., & Springer, M. (2007, Autumn). Teacher performance pay: A review. *Journal of Policy Analysis and Management, 26*(4)*,* 909-949. Retrieved from http://www.jstor.org/stable/30162809

Poltrock, L.A. (1984). Educational reform and its labor relations impact from a union perspective labor. *Journal of Law & Education, 13*(3), 457-476. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/jle13&id=467

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Reagan, R. (1985, February 28). Remarks at the national association of independent schools annual meeting. Retrieved November 1, 2012 from http://www.reagan.utexas.edu/archives/speeches/1985/22885a.htm

Riley, M. J. (1985-1986). The effect of present laws concerning tenure and certification upon proposed programs to create incentives for public school teachers toward excellence. *Southern University Law Review, 12*(2), 133-156. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/soulr12&id=139

Rivkin, S.G., Hanusheck, E.A., & Kane, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458. Retrieved from http://www.jstor.org/stable/3598793

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2), 247-252. Retrieved from http://www.jstor.org/stable/3592891

Rockoff, J.E., & Speroni, C. (2010). Subjective and objective evaluations of teacher effectiveness. *American Economic Review, 94*(2), 247-252. Retrieved from http://www.jstor.org/stable/27805001

Rockoff, J. E., & Speroni, C. (2011). Subjective and objective evaluations of teacher effectiveness: Evidence from New York City. *Labour Economics, 18*(5), 687-696. doi:10.1016/j.labeco.2011.02.004

Ross, S.A. (1973). The economic theory of agency: The principal's problem. *The American Economic Review, 63*(2), 134-139. Retrieved from: http://www.jstor.org/stable/1817064

Rothman, R. (2010). Beyond test scores: Adding value to assessment. *School Administrator, 67*(2), 20-24. Retrieved from http://libproxy.lib.unc.edu/login?url=http://search.proquest.com.libproxy.lib.unc.edu/docview/219248376?accountid=14244

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics, 125*(1), 175-214. doi: 10.1162/qjec.2010.125.1.175

Roye, N. G. (2010, Spring). High hopes hamstrung: How the "Trial de nuovo" for termination of tenured teachers' contracts undermines school reform in Oklahoma. *Oklahoma Law Review, 62*(3), 527. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/oklrv62&id=533

Sanders, W.J. (1998). Value-added assessment. *School Administrator, 55*(11), 24-27. Retrieved from http://search.proquest.com/docview/219243518?accountid=14244

Sanders, W.J., & Horn, S.P. (1994). The Tennessee Value-Added Assessment System: Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8,* 299-311. doi:10.1007/BF00973726

Sanders, W.J., & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *The Journal of Personnel Evaluation in Education, 12*(3), 247-256. Retrieved from http://dx.doi.org/10.1023/A%3A1008067210518

Sanders, W.L., & Rivers, J.C. (1996). Cumulative and residual effects of teachers on future student academic achievement. Research Progress Report. Knoxville: University of Tennessee Value-Added Research and Assessment Center

Sanders, W.L., Saxton, A.M., & Horn, S.P. (1997). The Tennessee Value-Added Assessment System (TVAAS): A quantitative, outcomes-based approach to educational assessment. In J. MIllman (Ed.), *Grading Teachers, Grading Schools*. Thousand Oaks, CA: Corwin Press.

Santelices, M. V., & Wilson, M. (2010). Unfair treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Educational Review, 80*(1), 106-133. Retrieved from http://libproxy.lib.unc.edu/login?url=https://search-proquest com.libproxy.lib.unc.edu/docview/212301931?accountid=14244

SAS. (2012). SAS EVAAS for K-12: Assess and predict student performance with precision and reliability. Retrieved from http://www.sas.com/govedu/edu/k12/evaas/index.html

S.C. Code Ann. § 59-18-1500 (2009)

Sharma, A., Joyner, A.M., & Osment, A. (2014). Adverse impact of racial isolation on student performance: A study in North Carolina. *Education Policy Analysis Archives, 22* (14). http://dx.doi.org/10.14507/epaa.v22n14.2014

Slaughter, J.E., Zickar, M.J., Highhouse, S., & Mohr, D.C. (2004). Personality trait inferences about organizations: Development of a measure and assessment of construct validity. *Journal of Applied Psychology, 89*(1), 85-103. doi:10.1037/0021-9010.89.1.85.

Spence, M. (1974). Job market signaling. *The Quarter Journal of Economics*, 87(3), 355-374. Retrieved from http://www.jstor.org/stable/1882010

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review, 92*(3), 434-459. Retrieved from http://www.jstor.org/stable/3083350

Springer, M.G. (2009). Rethinking teacher compensation policies: Why now, why again? In M.G. Springer (Ed.), *Performance incentives: Their growing impact on American K-12 Education*. Washington, DC: Brookings.

Springer, M.G., & Gardner, C.D. (2010). Teacher pay for performance: Context, status, and direction. *Phi Delta Kappa, 91(*8), 8-15. Retrieved from http://www.jstor.org/stable/20697111

Srivastava, J. (2001). The role of inferences in sequential bargaining with one-sided incomplete information: Some experimental evidence. *Organizational Behavior and Human Decision Processes, 85*(1), 166-187. doi:10.1006/obhd.2000.2936.

StataCorp. (2017). *Stata 15 base reference manual*. College Station, TX: Stata Press.

State of North Carolina, Office of the Governor. (2014). Governor McCrory announces sweeping array of education initiatives [press release]. Retrieved from

http://www.governor.state.nc.us/newsroom/press-releases/20140507/governor-mccrory-announces-sweeping-array-education-initiatives

Steele, C. M. (2003). Stereotype threat and African American student achievement. In T. Perry, C. Steele, & A. Hilliard, III (Eds.), *Young gifted and Black: Promoting high achievement among African-American students* (pp. 109-130). Boston, MA: Beacon.

Stewart, B. E. (2006). *Value-added modeling: The challenge of measuring educational outcomes.* New York: Carnegie Corporation.

Stiglitz, J.E. (2002). Information and the change in the paradigm in economics. *American Economic Review, 92*(3), 460-501. Retrieved from http://www.jstor.org/stable/3083351

Summers, A., & Wolfe, B. (1977). Do schools make a difference? *American Economic Review, 67*(4), 639-652. Retrieved from http://www.jstor.org/stable/1813396

Superfine, B. M. (2011, Winter). Stimulating school reform: The American recovery and reinvestment act and the shifting federal role in education. *Missouri Law Review, 76*(1), 81-134. Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA281113113&v=2.1&u=unc_main&it=r&p=&sw=w&asid=cde2ea4e3bd11e4ab51c511879244e84

Tex. Educ. Code Ann § 21.701-705 (2009)

The Friday Report. (2005, Feb 18). Retrieved December 5, 2014 from http://www.ncforum.org/doclib/2012_0615.pdf

Tucker, P.D. (1997). Lake Wobegon: where all teachers are competent (Or, have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education, 11*, 103-126. Retrieved from http://dx.doi.org/10.1023/A%3A1007962302463

Tyler, J.H., Taylor, E.S., Kane, T.J., & Wooten, A.L. (2010). Using student performance data to identify effective classroom practices. *The American Economic Review, 100*(2), 256-260. Retrieved from http://www.jstor.org/stable/27805000

U.S. Department of Education. (2006, May 17). *Secretary Spellings approves Tennessee and North Carolina growth model pilots for 2005-2006.* Retrieved from http://www.ed.gov/news/pressreleases/2006/05/05172006a.html

Vagi, R., Pivovarova, M., Barnard, W.M. (2017). Keeping our best? A survival analysis examining a measure of preservice teacher quality and teacher attrition. *Journal of Teacher Education*, 1-13, doi: 10.1177/0022487117725025.

Varma, A. & Stroh, L.K. (2001). The impact of same-sex LMX dyads on performance evaluations. *Human Resource Management, 40*(4), 309-320. doi:10.1002/hrm.1021

Veir, C. A., & Dagley, D. L. (2002). Legal issues in teacher evaluation legislation: A study of state statutory provisions. *Brigham Young University Education and Law Journal, 2002*, 1-16. Retrieved from http://heinonline.org.libproxy.lib.unc.edu/HOL/Page?handle=hein.journals/byuelj2002&collection=journals&page=1

Viadero, D. (2008). Scrutiny heightens for value added research methods. *Education Week, 27*(36), 1, 12-13. Retrieved from http://search.proquest.com/docview/202709515?accountid=14244

Wayne, S.J., & Ferris, G.R. (1990). Influence, tactics, affect, and exchange quality in supervisor-subordinate interactions: a laboratory experiment and field study. *The Journal of Applied Psychology, 75*(5), 487-499. Retrieved from https://auth-lib-unc-edu.libproxy.lib.unc.edu/ezproxy_auth.php?url=http://search.ebscohost.com.libproxy.lib.unc.edu/login.aspx?direct=true&db=pdh&AN=1991-08594-001&site=ehost-live&scope=site

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D. (2009). *The widget effect*. Brooklyn, New York: The New Teacher Project. Retrieved from https://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf.

Wright, S.P., Horn, S.P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement. *Journal of Personnel Evaluation in Education*, *11*, 57-67. Retrieved from http://dx.doi.org/10.1023/A%3A1007999204543

Zhang, Z., Verstegen, D.A., & Kim, H.R. (2008). Teacher compensation and school quality: New findings from national and international data, *Educational Considerations, 35*(2). Retrieved from https://auth-lib-unc-edu.libproxy.lib.unc.edu/ezproxy_auth.php?url=http://search.ebscohost.com.libproxy.lib.unc.edu/login.aspx?direct=true&db=eft&AN=508060947&site=ehost-live&scope=site

Zhang, Y., & Wiersema, M.F. (2009). Stock market reaction to CEO certification: The signaling role of CEO background. *Strategic Management Journal, 30*(7), 693-710. Retrieved by http://www.jstor.org/stable/20536072